Department of experimental and Health Sciences
Faculty of Health and Life Sciences
Universitat Pompeu Fabra
Doctoral Thesis 2015

Identification of new LOXL2 substrates

A role for lysine oxidation in embryonic stem cell differentiation

Dissertation presented by
Ane Iturbide Martinez de Albeniz
For the degree of Doctor in Biomedical Research

Work carried out under the supervision of Dr. Sandra Peiró and Dr. Antonio García de Herreros in the Chromatin Dynamics in Tumoral Cells Group in the Cancer Research Program in the Institut Hospital del Mar d'Investigacions Mèdiques (IMIM)

Barcelona, 2015

Dr. Sandra Peiró	Dr. Antonio García de Herreros	Ane Iturbide
(thesis co-director)	(thesis co-director)	(PhD student)

IMIM Parc de Salut MAR Barcelona
Institut Hospital del Mar d'Investigacions Mèdiques

upf. Universitat Pompeu Fabra Barcelona

Ene familiari
Borjari

Abstract

Protein function is often regulated and controlled by posttranslational modifications, such as oxidation. Although oxidation has been mainly considered to be uncontrolled and nonenzymatic, many enzymatic oxidations occur on enzyme-selected lysine residues; for instance, LOXL2 oxidizes lysines by converting the ε-amino groups into aldehyde groups. Using an unbiased proteomic approach, we have identified methylated TAF10, a member of the TFIID complex, as a LOXL2 substrate. LOXL2 oxidation of TAF10 induces its release from its promoters, leading to a block in TFIID-dependent gene transcription. In embryonic stem cells, this results in the inactivation of the pluripotency genes and loss of the pluripotent capacity. During zebrafish development, the absence of LOXL2 resulted in the aberrant overexpression of the neural progenitor gene Sox2 and impaired neural differentiation. Thus, lysine oxidation of the transcription factor TAF10 is a controlled protein modification and demonstrates a role for protein oxidation in regulating pluripotency genes.

La funció de les proteïnes és sovint regulada i controlada per modificacions post-traduccionals, com ara l'oxidació. Encara que l'oxidació ha sigut principalment considerada com una modificació incontrolada i no enzimàtica, moltes oxidacions tenen lloc en residus de lisina seleccionats enzimàticament: per exemple, LOXL2 oxida lisines convertint els grups ε-amino en grups aldehid. Mitjançant espectrometria de masses hem identificat el TAF10 metilat (membre del complex TFIID) com a substrat de LOXL2. L'oxidació de LOXL2 sobre TAF10 indueix la seva separació dels promotors, donant lloc a un bloqueig de la transcripció dels gens dependents de TFIID. En cèl·lules mare embrionàries, aquest efecte resulta en la inactivació dels gens de pluripotència i en la pèrdua de la capacitat pluripotent. La importància de la LOXL2 en la pèrdua d'aquest estat pluripotent queda reflexada en el desenvolupament del peix zebra, on la seva absència resulta en una sobreexpressió del gen marcador de progenitors neurals Sox2 i en la deterioració de la diferenciació neural. Per tant, l'oxidació de la lisina del factor de transcripció TAF10 és una modificació proteica controlada i demostra com l'oxidació de proteïnes pot exercir una funció en la regulació dels gens de pluripotència.

Table of contents

Figure index

Introduction

Results

Concluding remarks

Table index

Introduction

Materials and methods

Annex

Acronyms and abbreviations

2i: inhibitors CHIR99021 and PD0325901
3C: chromosome conformation capture
53BP1: p53-binding protein 1
AP: alkaline phosphatase
ATM: serine-protein kinase ATM
ATRA: *all-trans* retinoic acid
Bmp: bone morphogenetic protein
c-Myc: V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CAF-1: chromatin assembly factor 1
CCNE1: Cyclin E1
ChIP: chromatin immunoprecipitation
CNS: central nervous system
CpG: cytosine-guanine dinucleotide
CRABP1-2: cellular retinoic acid binding proteins 1-2
CSC: cancer stem cell
CYP26: Cytochrome P450 26A1
DCE: downstream core element
DDR: DNA damage response
DNA: deoxyribonucleic acid
DNMT: DNA methyltransferase
DPE: downstream promoter element
DUB: deubiquitylases
EB: embryonic body
ECM: extracellular matrix
EMT: epithelial-to-mesenchymal transition
ERF1: eukaryotic translation termination factor 1
ESC: embryonic stem cell
Esrrb: estrogen-related receptor beta
FAD: flavin adenine dinucleotide
FBRL: rRNA 2'-O-methyltransferase fibrillarin
FGF: fibroblast growth factor
FoxD3-4: forkhead Box D3-4
G6PD: glucose-6-phosphate 1-dehydrogenase
GCN5: histone acetyltransferase GCN5
GCNF: nuclear receptor subfamily 6 group A member 1
GO: gene ontology
GPI: glucose-6-phosphate isomerase
Gsk3β: glycogen synthase kinase-3 β
GTF: general transcription factor
H: histone
HA: hemagglutinin
HAT: histone acetyltransferase
HDAC: histone deacetylase
hDPSC: human dental pulp stem cell
HFD: histone fold domain
Hi-C: high-throughput sequencing of chromosome conformation capture
HNSCC: head and neck squamous cell carcinoma
HoxA1-A7: homeobox A1-7
HP1: heterochromatin protein 1
HPRT: hypoxanthine phosphoribosyltransferase
HRE: hormone response element
ICM: inner cell mass
IP: immunoprecipitation
iPS: induced pluripotent stem
JmjC: jumonji C
KDM: lysine demethylase
KDM1A: lysine (K)-specific demethylase 1A
KEGG: Kyoto encyclopedia of genes and genomes
Klf2-5: Kruppel-like Factor 2-5
KMT: lysine methyltransferase
KMT7: SET domain containing (lysine methyltransferase) 7
LIF: leukaemia inhibitory factor
LOX: lysyl oxidase
LOXL1-4: lysyl oxidase-like 1-4
LSD1: lysine (K)-specific demethylase 1A
LTQ: lysyl tyrosylquinone

MAPK: mitogen-activated protein kinase
MET: mesenchymal-to-epithelial transition
MG132: carbobenzoxy-Leu-Leu-leucinal
MNase: micrococcal nuclease
MRE11: meiotic recombination 11 homolog A
MRN: Mre11-Rad50-Nbs1 complex
mRNA: messenger RNA
MS: mass spectrometry
NADPH: nicotinamide adenine dinucleotide phosphate
Nanog: nanog homeobox
NBS1: Nijmegen Breakage syndrome 1
NeuroD: neuronal differentiation 1
NeuroG1: neurogenin1
NF: neurofibromin
NFkB: nuclear factor kappa-B
Notch1: notch homolog 1
Oct3/4: POU class 5 homeobox 1
Otx2: orthodenticle homeobox 2
p53: tumor protein p53
PAD: protein arginine deiminase
PARP3: poly (ADP-ribose) polymerase family, member 3
Pax6: paired box 6
PcG: polycomb group
PCR: polymerase chain reaction
PE: primitive endoderm
PEI: polyethylenimine polymer
PIC: pre-initiation complex
PLK1: polo-like kinase 1
polII: RNA polymerase II
Pou5f1: POU class 5 homeobox 1
PPAR: peroxisome proliferator-activated receptors
PRC1-2: polycomb repressive complex 1-2
PS: primitive streak
PTEN: phosphatase and tensin homolog
PTM: posttranslational modification
qRT-PCR: real-time polymerase chain reaction
RA: retinoic acid
RALDH: retinaldehyde dehydrogenases
RAR: retinoica cid receptor
RARE: retinoic acid response element
RB: retinoblastoma 1
RBP1/4: retinol-binding protein 1/4
Rex1: zinc finger protein 42
RNA: ribonucleic acid
RNAPII: RNA polymerase II
RXR: retinoic X receptor
SAGA: Spt-Ada-Gcn5-Acetyl transferase
Sall4: spalt-like transcription factor 4
SET7/9: histone-lysine N-methyltransferase SETD7
SETD8: N-lysine methyltransferase SETD8
shRNA: short hairpin RNA
SMAD: mothers against decapentaplegic homolog 1
SNAIL1: zinc finger protein SNAI1
Sox1-2: SRY (sex determining region Y)-b box1-2
SRCR: scavenger receptor cysteine-rich
STAT3: signal transducer and activator of transcription 3
STRA6: stimulated by retinoic acid 6
SUV39H1: suppressor of variegation 3-9 homolog 1
TAD: topologically associated domain
TAF1-13: TBP associated factor 1-13
TALDO: transaldolase
TBP: TATA-box-binding protein
Tbx3: T-box protein 3
TE: trophectoderm
TF: transcription factor
Tfcp2l1: transcription factor CP2-like 1
TFTC: TATA-binding protein-free TAF-containing complex
TGF-β: transforming growth factor beta
Tip60: histone acetyltransferase KAT5
TKT: transketolase
TSS: transcription start site

Tuj1: neuron-specific class III beta-tubulin

TWIST1: twist Family BHLH transcription factor 1

ZEB1/2: zinc finger E-Box binding homeobox 1/2

Zic1: Zic family member 1

INTRODUCTION

1. Chromatin organization and gene transcription

1.1. Chromatin structure

Mammalian cells pack 1.7 metres of DNA into a 5-micrometre nucleus in a form that allows it to be replicated and transcribed in stable, tissue-specific patterns. For this, DNA is highly condensed in a nucleoprotein complex named chromatin. The basic unit of chromatin is the nucleosome, each of which wraps 147 bp of DNA in two turns around a histone octamer[1]. These octamers are composed by two H3-H4 and two H2A-H2B dimers, and an additional histone H1 binds to the core particles and protects the internucleosomal linker DNA (*Figure I.1*).

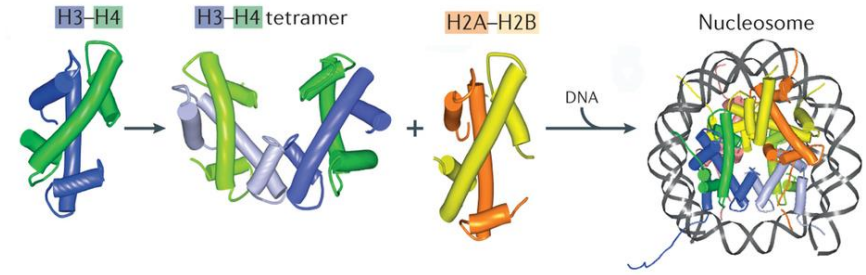

Figure I.1. Crystal structure of the nucleosome. The nucleosome consists of 147 bp of DNA (strand colored grey) wrapped 1.7 times around a core octamer of histone proteins with two H2A (orange) H2B (yellow) dimers associated with a H3 (blue) H4 (green) tetramer[2].

Further condensation is obtained by supercoiling DNA. This process allows organization of DNA and regulation of transcription[3]. At the kilobase to megabase scale, distal

regulatory elements come into direct contact with their targets via chromatin loops.[4] At the megabase scale, co-occupation of functional sites have been observed for some genes such as foci of Polycomb proteins[5]. At the nuclei scale, chromosomes are located in discrete territories, organized to place gene-poor chromosomes in the periphery and gene-rich regions in the interior[6].

Over the past few years, transcriptional activity of some genes has been correlated with their nuclear positioning, depending on their proximity to the repressive nuclear lamina of the periphery or their position in the bulk of the chromosome territory[7,8]. Interestingly, it has been recently shown that chromatin decondensation and transcriptional activation are not always related, as opening chromatin without gene activation is sufficient for relocalization of genes to the nuclear interior[9].

Chromatin conformation has been recently studied by combining Chromosome Conformation Capture (3C) technology[10] to high-throughput sequencing (Hi-C)[11]. These studies first demonstrated that active chromatin predominantly associates with other active regions, and repressed chromatin associates with other silent regions with little inter-mixing of the two types. More recently, high-resolution chromatin interaction maps revealed that genomes fold into distinct modules called topologically associated

domains (TADs) (*Figure I.2*). Genomic interactions are strong inside these domains but are sharply depleted on crossing the boundary between two TADs[12,13,14]. These domains correlate with markers of chromatin activity (such as histone modifications and replication timing) and contain coordinately regulated genes.

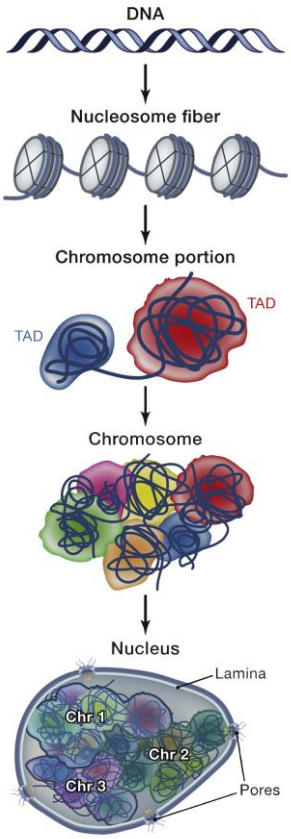

Figure I.2. Analogous hierarchical organization of genome structure. Primary structure comprising the nucleotide sequence (packaged into a nucleosomal fibre in eukaryotic chromatin) on a single polymeric chain form locally stabilized interactions to fold into secondary structures, chromatin TADs. These domains in turn hierarchically co-associate to form a tertiary structure of a chromosome. The co-associations of multiple, separately encoded subunits, form the final quaternary structure of an entire genome[15].

Higher-order chromosomal structures may be built up from key stabilizing interactions between regulatory elements inside TADs[16]. However, genome folding is not completely defined, probably because there is not a specific final conformation since genome structure is highly dynamic. When performing single-cell experiments, high variability of genomic configurations have been observed, reinforcing this idea[17].

1.1.1. *Chromatin classification into euchromatin and heterochromatin*

Nearly 100 years ago in 1928, heterochromatin was first distinguished from euchromatin on the basis of differential compaction at interphase[18]. The initial classification of DNA was based on the observation that euchromatic regions change their degree of condensation during cell cycle division, whereas heterochromatic regions remain highly condensed throughout most of the cell cycle.

Euchromatin is less condensed, more accessible and generally more easily transcribed, whereas heterochromatin is typically highly condensed, inaccessible and highly ordered in nucleosomal arrays[19]. The biological significance of heterochromatin remained obscure for many years, since RNA-DNA hybridization experiments suggested that these

regions were transcriptionally silent[20]. However, with increased sensitivity of molecular techniques, transcription of centromeric and pericentromeric regions was confirmed[21]. It is now apparent that heterochromatin plays a number of biological roles, including repression of transcription and recombination[22] and mediation of proper chromosome segregation[23,24] and long-range chromatin interactions[25].

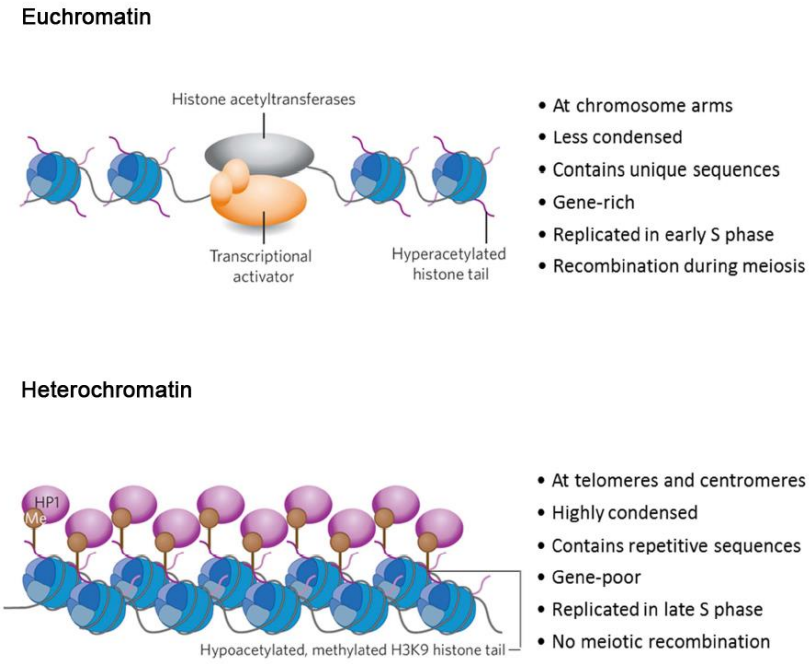

Figure I.3. Properties of euchromatic and heterochromatic regions. Main characteristic of each chromatin subtype are listed. However, since heterochromatin is very difficult to describe and most of the shown properties have exceptions, characteristics shown in this case can be more clearly related to pericentromeric heterochromatin[26].

Different roles for heterochromatic RNAs have also been described. For example, RNA derived from repeat sequences is related to general assembly of heterochromatin, since HP1 requires RNA to assemble condensed chromatin in mammalian cells[27]. Other studies have also correlated transcriptional activation of pericentromeric regions with decondensation of heterochromatin that goes along with an increase in active histone marks[28].

Heterochromatin can be divided into two classes: 1) constitutive heterochromatin, which includes chromosomal regions that contain a high density of repetitive DNA elements, such as clusters of satellite sequences and transposable elements. These are found at centromeres and telomeres and remain condensed throughout the cell cycle; and 2) facultative heterochromatin, which includes developmentally regulated loci, at which the chromatin stage can change in response to cellular signals and gene activity.

Heterochromatin can propagate and influence gene expression in a region-specific but sequence-independent manner. Heterochromatin spreading across domain causes epigenetic repression and transcription silencing. However, several works have reported that heterochromatin formation is required to activate gene transcription[29,30]. Hence, it seems that heterochromatin can act as a platform for recruiting both silencing and anti-silencing factors.

1.2. Epigenetics

Epigenetics refers to those processes that ensure the inheritance of variation ("genetic") above and beyond ("epi-") changes in DNA sequence. This information resides in self-propagating molecular signatures that provide a memory of previously experienced stimuli, without irreversible changes in the genetic information[31]. Epigenetic modifications are mainly related to DNA methylation and histone modifications.

1.2.1. Histone modifications

Both histone tails and globular domains are subject to a vast array of posttranslational modifications (PTMs), summarized in *Figure I.4*. Some of these modifications are associated with active transcription, such as acetylation of histone H3 and histone H4 or di- or trimethylation of lysine 4 in histone H3 (H3K4). Others are related to inactive genes or regions, such as H3K9 methylation or H3K27 methylation.

Histone modifications exert their effects via two main mechanisms: by influencing the overall structure of chromatin, and by binding effector molecules. Modifications affecting higher-order chromatin structure seem to work by affecting the contact between different histones in adjacent nucleosomes or directly interaction between histones and

DNA. Acetylation is the best-known modification to unfold chromatin since it neutralizes the basic charge of the lysine residues[32]. In general, any alteration in histone charge could have an effect on nucleosome organization.

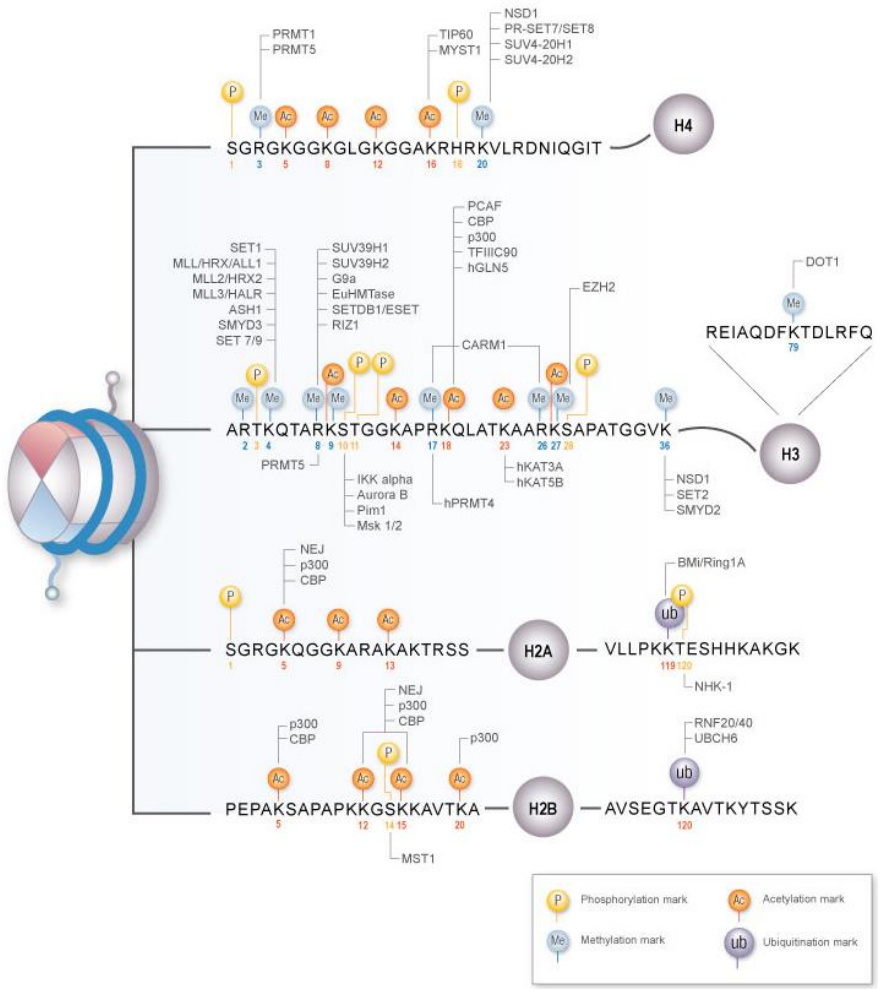

Figure I.4. Map of histone modifications. Sequence of the four human core histones with main posttranslational modifications and modifying enzymes indicated. The modifications include acetylation (Ac), methylation (Me), phosphorylation (P) and ubiquitination (ub). Adapted from epitomics website.

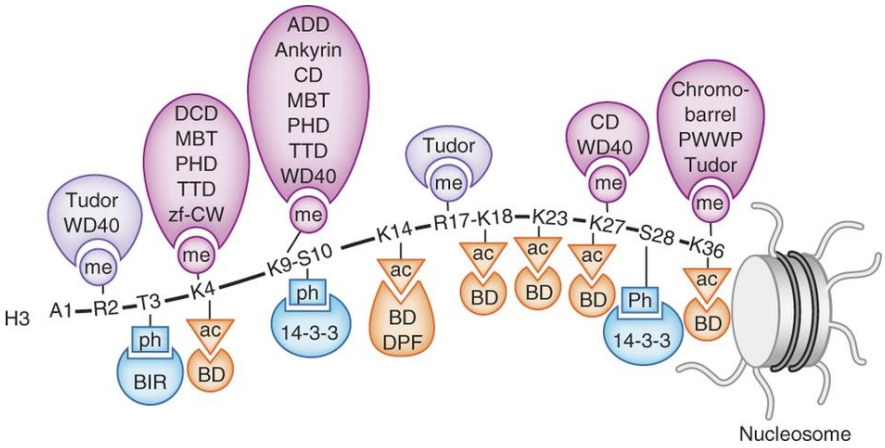

Figure I.5. Readers of histone posttranslational modifications. Posttranslational modifications in histone H3 N-terminal tail, comprising methylation (me), phosphorylation (ph) and acetylation (ac), are recognized by different protein domains within histone readers[33].

The function of effector molecules is a better-studied effect of histone modifications. Different histone modifications are recognized by specific domains in effector proteins (*Figure I.5*) with the objective of recruiting enzymatic activities onto chromatin. These activities are related not only to regulating gene expression[34,35,36] but also to generating DNA repair foci[37,38], DNA replication signalling[39,40] and chromosome condensation[41]. The term “histone code” has been loosely used to describe the role of modifications to enable DNA functions. However, it is unlikely the existence of a predictable “code” related to histone modifications[42].

1.2.1.1. Histone modifying enzymes

Each of the epigenetic pathways mentioned above requires enzymes that transfer those modifications, known as the writers, and enzymes to revert the modifications, known as editors or erasers. These enzymes are also in general referred to as histone modifying enzymes.

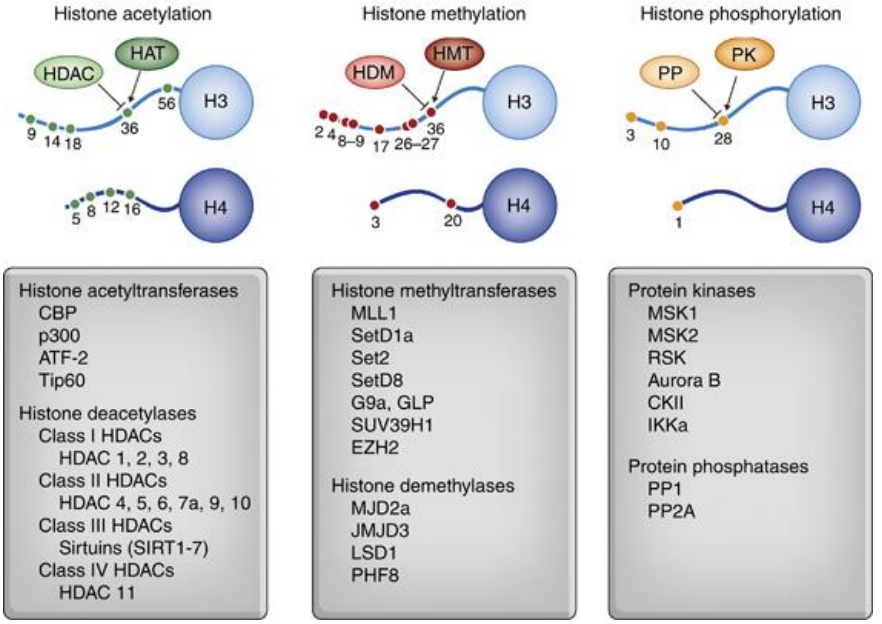

Figure I.6. Summary of most studied histone modifying enzymes. Main histone modifications and related writer and editor enzyme families are represented. Examples of different enzymes for each family are listed below[43].

Histone modifying enzymes are usually specific to particular amino acid motifs and can target both core histones and histone variants. Although a wide range of PTM types have been described so far, the best studied ones include histone

acetyltransferases (HATs), histone deacetylases (HDACs), histone kinases, histone phosphatases, lysine methyltransferases (KMTs), lysine demethylases (KDMs), ubiquitylation enzymes (E1, E2 and E3 enzymes) and deubiquitylases (DUBs). These enzymes often exist in multi-subunit complexes and modify specific residues either on the amino-terminal tails or within the globular domains of histones (*Figure I.6*). For example, in the two repressive Polycomb group (PcG) protein complexes, Polycomb repressive complex 1 (PRC1) contains either RING1A or RING1B, both of which catalyse the monoubiquitylation of histone H2A at lysine 119, and PRC2 contains EZH2 enzyme, which catalyses the trimethylation of H3K27[44].

It should be noted that the cellular enzymes that modify histones may also have non-histone targets and, as such, it is difficult to divorce the cellular consequences of individual modifications from the boarder targets of many of these enzymes. PTMs on non-histone proteins can regulate protein-protein interactions, stability, localization, and/or enzymatic activities of proteins involved in diverse cellular processes. For example, histone acetyltransferases activity has been shown to contribute to malignant transformation through altered global histone acetylation patterns. However, it is also well established that several non-histone proteins, including many important oncogenes and tumour suppressors such as MYC, p53, and PTEN, are also dynamically acetylated[45].

Another example relies on lysine methylation. Although nearly 80 enzymes have been shown to dynamically regulate histone lysine methylation, few non-histone proteins have been also reported as substrates of these enzymes. Methylation of p53 protein by KMT7/SET7 histone methyltransferase was the first reported KMT-mediated methylation event on a non-histone protein[46]. Since that discovery, several histone methyltransferases and demethylases have been identified as p53 regulators[47,48], as well as other non-histone substrates, such as NFκB[49], STAT3[50] and RB[51].

1.2.2. DNA methylation

DNA methylation is one of the best-characterized chemical modifications of chromatin, described in both prokaryotes and eukaryotes. In prokaryotes, it is used as a defence against bacteriophages and occurs at both cytosine and adenine residues[52]. However, in multicellular organisms, DNA methylation is found almost exclusively at cytosine residues and contributes to epigenetic regulation of gene expression by reducing the transcriptional activity of chromatin[53]. In mammals, nearly all DNA methylation occurs on cytosine residues of CpG dinucleotides. In the human genome, 60–80% of the 28 million CpG dinucleotides are methylated[54]. Regions of the genome containing high CpGs density are

referred to as CpG islands and are positioned at the 5′ ends of many human genes. In fact, it is estimated that around 60% of human genes are associated with CpG islands, of which the great majority are unmethylated[55].

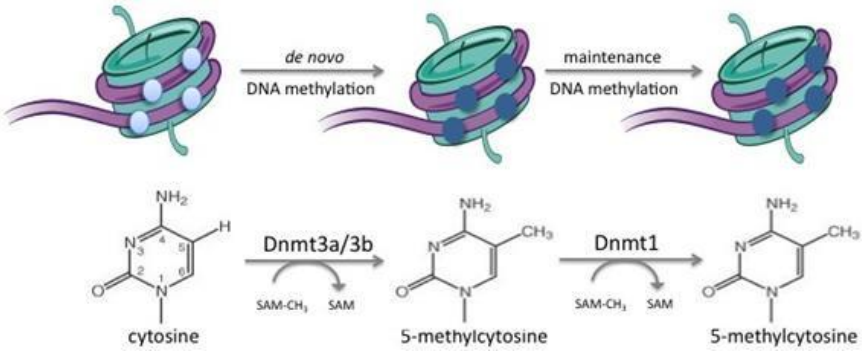

Figure I.7. DNA methyltransferases. Representation of DNA methylation by *de novo* and maintenance methyltransferases: DNMT3a/3b methylates the 5′ position of cytosine *de novo*. Following DNA replication, DNMT1 maintains the pattern of DNA methylation. Adapted from Meissner laboratory's website.

Cytosine methylation can be donated by *de novo* or maintenance DNA methyltransferases (DNMTs). *De novo* DNMTs act after DNA replication, whereas maintenance methyltransferases add methyl groups to DNA during replication (*Figure I.7*). This is the case of DNMT1, which re-establishes symmetrical CpG methylation on newly synthesized hemi-methylated DNA[56]. *De novo* methylation occurs mostly in germ cells or the early embryo[57] and in adult somatic cells in certain tissues during aging[58]. However, it can also occur in pathological situations such as cancer. DNA methylation plays a role in many cellular processes, such as

repetitive and centromeric sequences silencing[59], mammalian imprinting[60] and X chromosome inactivation[61].

1.3. Transcriptional regulation

A variety and number of genes are transcribed into protein coding and noncoding RNA in mammalian cells. Transcriptional regulation of gene expression occurs thanks to the actions of a diverse range of factors, including chromatin remodellers, transcription factors, polymerases, helicases, topoisomerases and histone modifying enzymes, among others. Gene transcription regulation is mostly described as binding of transcription factors to enhancer elements and recruiting cofactors and RNA polymerase II to target genes.[62] However, multiple models of transcription regulation have been proposed in the past decades.

The canonical view of gene regulation has been based on sequential and ordered recruitment of factors. In this model, transcriptional activation is the result of a series of events that occur in a certain sequence. Nowadays, however, most models agree with a probabilistic model, in which transcription regulation is represented as an equilibrium thermodynamic phenomenon. In this case, stochastic interactions occur between transcription factors and DNA, which can account for the timing of the downstream transcriptional output resulting

from those interactions[63]. In fact, average occupancy of a binding site at a promoter by a transcription factor does not correlate nearly as well with expression levels as the occupancy time of that factor[64].

1.3.1. Transcription initiation

Transcription can be divided into two basic stages: initiation and elongation. In the first step of transcription initiation, an activator molecule binds to a recognition site upstream of the transcription start site (TSS) and recruits co-activators and chromatin remodelling machinery that leads to a nucleosome displacement at the core of the promoter, uncovering binding sites for the core machinery[65]. Thenceforth, the pre-initiation complex (PIC) assembles in an ordered manner and promotes RNA synthesis (*Figure I.8)*. In the second stage, the polymerase tracks along the DNA and makes an RNA copy[66].

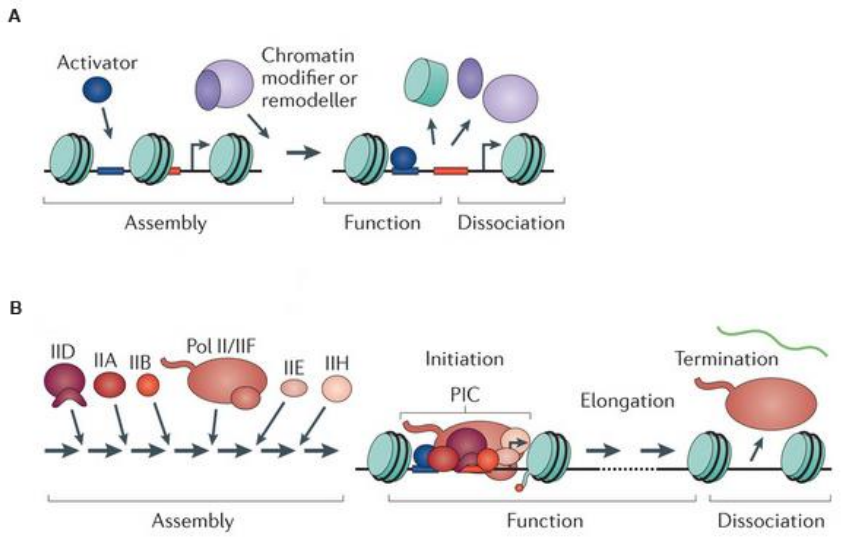

Figure I.8. The assembly-function-dissociation model. Representation of different stages and molecular players involved in transcription initiation. (**A**) Uncovering of binding sites for the core machinery by activators and chromatin remodellers. (**B**) Ordered assembly of the pre-initiation complex (PIC), transcription initiation, elongation and termination[63].

When gene transcription was first reproduced *in vitro*, it was shown that RNA polymerase by itself is not capable of specific initiation[67]. However, when additional protein complexes were added, selective transcription initiation was reproduced[68]. These protein factors are referred to as general transcription factors (GTFs).

1.3.2. General transcription factors

When first purified together with RNA polymerase II, four enzymatically active fractions (A, B, C and D) were detected,

which corresponded to the sequentially-eluted nuclear proteins. The protein factors presented in the A and D fractions were named TFIIA and TFIID, while the C fraction was fractionated into accessory factors TFIIB, TFIIE, TFIIF and TFIIH[69]. These make up the GTFs, which were named based on their purification using the following nomenclature: TF, Transcription Factor; the Roman numeral II, polymerase II–driven transcription; and the letter, the chromatographic fraction from which the specific GTF was isolated.

GTF assembly at the promoter is hierarchical. TFIID first recognizes the TATA box, followed by the entry of TFIIA, TFIIB, which help to stabilize promoter-bound TFIID, and then the recruitment of polII/TFIIF. After the formation of the stable TFIID-TFIIA-TFIIB-polII/TFIIF-promoter complex, TFIIE is recruited, and finally, TFIIH enters[70]. This process is known as the sequential pathway of PIC assembly, which collectively specifies the transcription start site.

1.3.2.1. TFIID

TFIID is a multiprotein complex comprising TBP and 13 TBP-associated factors (TAFs), with molecular weights ranging from 15 kDa to 250 kDa (*Figure I.9*). Structure analysis revealed that there is a striking similarity in the amino sequences of TAF6, TAF9 and TAF12 with the core histones

H4, H3 and H2B, respectively. Moreover, nine TAFs have histone fold domains, which function as a building block of the complex[71].

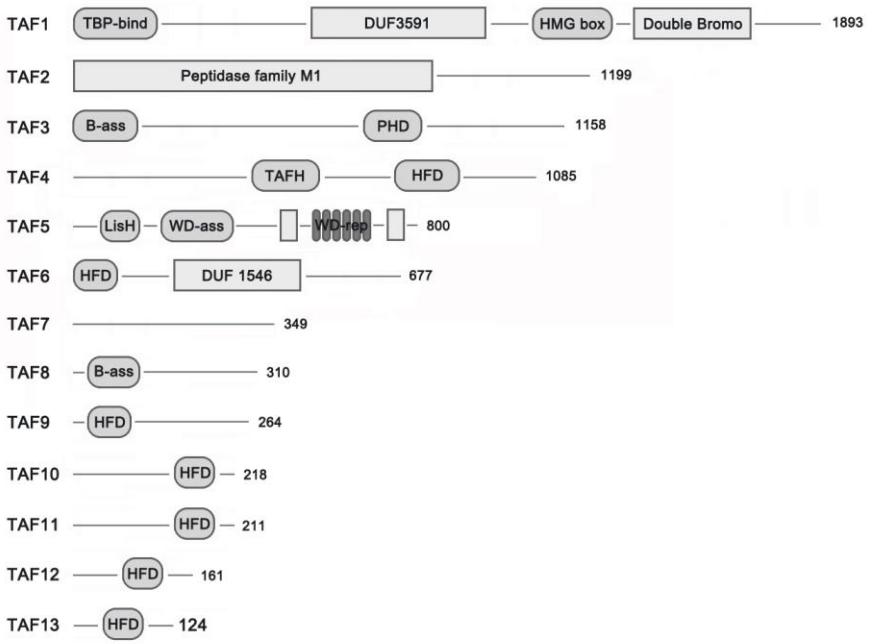

Figure I.9 Schematic representation of the TAF family members. TBP-bind, TATA box-binding protein binding; DUF3591, domain of unknown function 3591; Double Bromo, double bromodomain; B-ass, bromodomain associated; PHD, PHD-finger; TAFH, TAF homology; HFD, histone fold domain; LisH, LIS1 homology; DUF1546, domain of unknown function 1546[72].

The TFIID complex is able to bind core promoter elements through several components. For example, TBP is able to bind the TATA box (an A/T-rich sequence located approximately 25–30 nucleotides upstream of the TSS), while other TAFs are able to bind Inr (a pyrimidine-rich sequence

surrounding TSS), downstream promoter element (DPE) and downstream core element (DCE), giving the complex the capacity to bind both TATA-containing and TATA-less promoters[73].

Besides core promoter elements, histone H3K4 trimethylation has also been implicated in binding TFIID to the promoter region. This interaction is mediated by H3K4me3 recognition by the TAF3 plant homeodomain[74]. Moreover, H3K9ac and H3K14ac are recognized by TAF1 through its bromodomain[75].

The TFIID architecture was recently described to comprise three distinct subassemblies[76]. A symmetric core is composed by TAF4, TAF5, TAF6, TAF9 and TAF12, which then binds a TAF8-TAF10 dimer, resulting in an asymmetric structure that serves as a scaffold to nucleate the final holo-TFIID assembly by binding one copy of each of the remaining TAFs and TBP (*Figure I.10*).

Interestingly, a TFIID subcomplex composed by TAF2, TAF8 and TAF10 that assembles in the cytoplasm has also been identified[77]. It is based on the interaction of TAF8 and TAF10 histone fold domains, which bind to TAF2 to incorporate it into the core-TFIID complex. In this case, a TAF8-TAF10 dimer may function like a chaperone to regulate nuclear import and integration of TAF2 into core-TFIID.

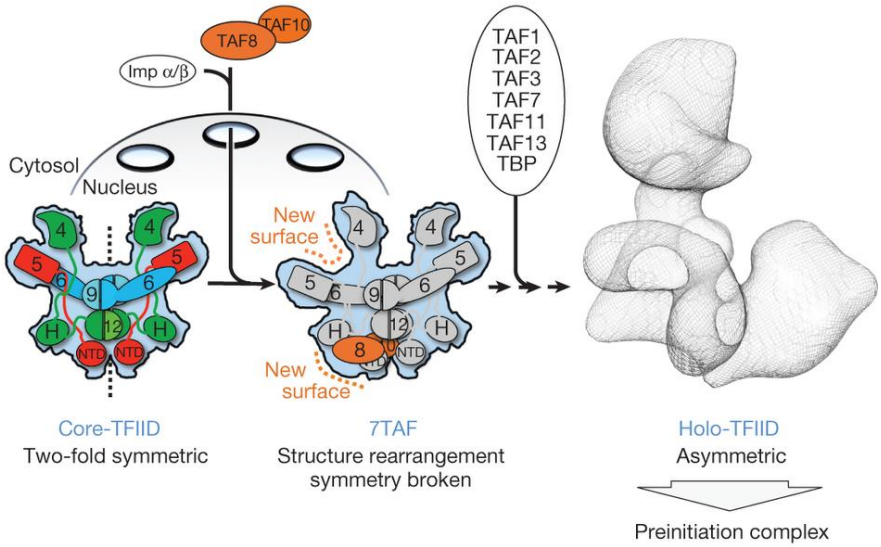

Figure I.10. Model for holo-TFIID assembly. Core-TFIID with two copies of TAF4, TAF5, TAF6, TAF9 and TAF12 is symmetric (left). The binding of the TAF8-TAF10 complex (orange) breaks the symmetry in core-TFIID, resulting in asymmetric 7TAF complex (middle). 7TAF exhibits two distinct halves and new binding surfaces for further subunits (dashed lines). Assembling with a single copy of the remaining TAFs and TBP results in asymmetric clamp-shaped holo-TFIID (grey mesh) that nucleates the pre-initiation complex[76].

TFIID subunits and TBP paralogues have unique functions during development, differentiation and cell proliferation. For example, TAF8 is upregulated during adipogenesis[78], while TAF10 has been related to liver development[79]. Moreover, mouse embryonic stem cells (ESCs) express high levels of TAF5 and TAF6 subunits as compared to other cell lines. Concomitantly, the TFIID complex was demonstrated to be required to maintain the pluripotent circuitry in ESCs, as high levels of expression are necessary for active transcription of Nanog, Oct4 and Sox2 pluripotency genes[80].

2. Lysyl oxidase-like 2 (LOXL2)

2.1. LOX family of proteins

The lysyl oxidase family of proteins comprises the lysyl oxidase (LOX) and four lysyl oxidase–like proteins (LOXL1–4). These proteins are amine oxidase enzymes able to remove the amino group located in the ε-position of lysine residues through oxidation, leaving an aldehyde group as product of the reaction. They depend on two cofactors to perform the reaction, copper and quinone[81].

The LOX protein was first known to crosslink collagen and elastin in the extracellular matrix (ECM)[82]. However, recent research on this protein family has revealed several new functions beyond ECM organization, including tumour suppression[83], chemotaxis[84,85] and hypoxia-induced metastasis[86], that are related to intracellular functions of the enzymes.

The structure of the LOX protein family is based on a conserved C-terminal region that contains all the elements required for the enzymatic activity, including a copper-binding motif, residues for lysine tyrosylquinone cofactor formation and a cytokine receptor-like domain, and an N-terminal domain that gives sequence variability to each member (*Figure I.11*).

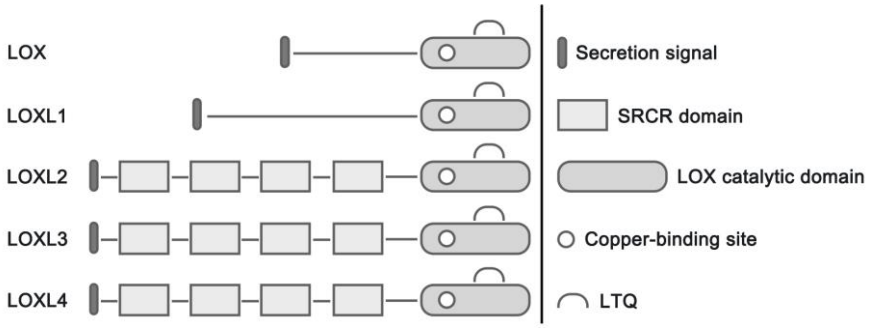

Figure I.11. Schematic representation of the LOX family members. Different domains of LOX proteins are represented. SRCR stands for scavenger receptor cysteine-rich, and LTQ, for lysyl tyrosylquinone[87].

Nuclear-related functions have been described for some of the members of this family. LOX protein is known to cause chromatin decondensation, together with mitotic abnormalities, micronuclei appearance and cell death when overexpressed[88]. This function has been related to a possible oxidation of histone H1 by LOX protein, the reaction may promote a loss of positive charge in histone H1 and favour detachment from DNA[89].

LOXL2 has also been related to nuclear functions. It is able to interact with several co-repressors, and its intracellular expression patterns have been related to cancer and increased malignancy, topics that will be discussed in later chapters.

2.2. LOXL2 as histone modifying enzyme

A histone modifying activity has been recently described for LOXL2. In contrast to the LOX protein, this function has been well characterized. LOXL2 specifically removes the amino group from trimethylated lysine 4 in histone H3. Since H3K4me3 is a mark that induces transcription, LOXL2 acts as a transcription repressor. This function has demonstrated to be key in E-cadherin repression, as LOXL2 is recruited together with SNAIL transcription factor to the promoter of the gene during epithelial-to-mesenchymal transition (EMT), where LOXL2 deaminates histone H3[90,91].

Moreover, LOXL2 oxidation of histone H3 has also been shown to be important for the regulation of heterochromatin transcription during EMT[92]. LOXL2, together with SNAIL, downregulates major satellite transcription from pericentromeric regions through H3K4me3 deamination. Subsequently, HP1α is released from pericentromeric regions, a step required for proper chromatin reorganization and EMT completion.

LOXL2 activity as a histone-modifying enzyme has also been observed in premalignant lesions in a human head and neck squamous cell carcinoma (HNSCC) model, where LOXL2 negatively regulates epidermal differentiation and the Notch1 signalling pathway. In this case, LOXL2 binds to at least two

different regions of the NOTCH1 promoter and reduces methylation levels of H3K4me3 and the subsequent RNA polymerase II recruitment[93].

2.3. LOXL2 in cancer

LOXL2 has been proposed to promote tumour cell survival, chemoresistance, regulates cell adhesion, motility and invasion, and remodels the tumour microenvironment. In fact, upregulation of LOXL2 has been observed in a number of human cancers[94,95] and seems to correlate with tumour grade, poor prognosis and decreased survival[96]. New studies even suggest a possible role for LOXL2 in pre-metastatic niche formation[97].

2.3.1. LOXL2 and EMT

A role for LOXL2 in EMT has also been characterized. EMT refers to the transdifferentiation of epithelial cells into motile mesenchymal cells, which is integral in development[98], wound healing[99] and stem cell behaviour[100], and which also contributes pathologically to fibrosis[101] and cancer progression[99]. During EMT, epithelial cells lose their junctions and apical-basal polarity, reorganize their cytoskeleton and undergo a change in the signalling programmes that define

cells and enable the development of an invasive phenotype. The reverse process is termed mesenchymal-to-epithelial transition (MET); interestingly, plasticity of the epithelial phenotype enables cells to transition through multiple rounds of EMT and MET.

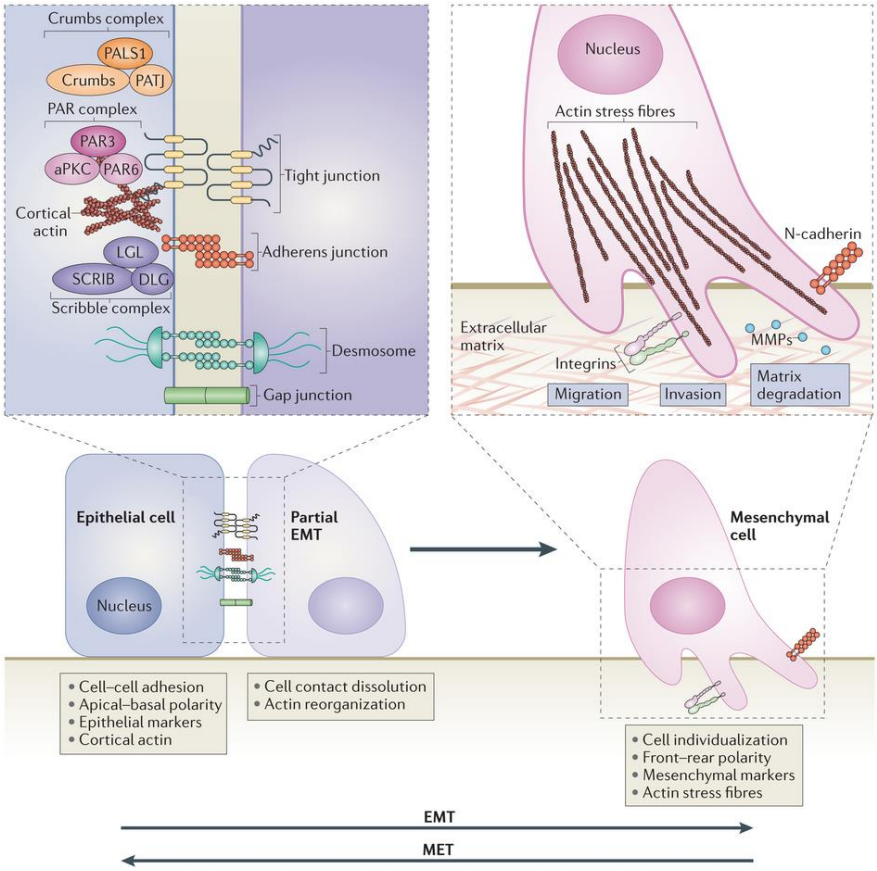

Figure I.12. Cellular events during EMT. Representation of main steps of EMT. First, there is a disassembly of epithelial cell-cell contacts and the loss of cell polarity and the expression of epithelial genes is repressed, concomitantly with the activation of mesenchymal gene expression. Next, the actin architecture reorganizes, and cells acquire motility and invasive capacities[102].

Upon the initiation of EMT, cell-cell junctions are deconstructed, and the junction proteins are relocalized and/or degraded. As the process progresses, the junction proteins themselves are transcriptionally repressed, such as for E-cadherin, which results in stable loss of epithelial junctions[103]. Cell also redirect their gene expression programme to produce changes in cytoskeletal architecture, to promote adhesion to mesenchymal cells and to alter interaction of cells with ECM. For example, the intermediate filament composition changes with the expression of cytokeratin and the activation of vimentin expression, enabling cell motility[104].

This switch in cell differentiation and behaviour is mediated by key transcription factors, including SNAIL1, TWIST1 and zinc-finger E-box-binding (ZEB1/2), which are tightly regulated at transcriptional, translational and posttranslational levels. Their expression is activated in early EMT, but all have different expression profiles, and their contributions to EMT depend on the cell or tissue type involved and the signalling pathways that initiate EMT. They often control expression of each other and functionally cooperate at target genes[105]. Together, these master regulators coordinate repression of epithelial genes and the induction of mesenchymal genes, and frequently the same transcription factors direct both repression and activation[103].

As mentioned above, both EMT and MET have been closely linked to stemness properties in development and cancer. The pluripotent ESCs in the inner cell mass of the blastocyst have epithelial characteristics and, during gastrulation, undergo the EMT process and form the primary mesoderm[106,107]. Conversely, reprogramming fibroblasts into induced pluripotent stem (iPS) cells requires the MET process, as it involves the repression of mesenchymal genes, including some that encode transcription factors with a role in EMT, and the activation of epithelial genes encoding epithelial cell junction proteins[108].

Due to this clear relation with stem cell properties, EMT has also been associated to carcinoma stem cell properties. For example, expression of SNAIL1 or TWIST1 in mammary epithelial cells induces a mesenchymal cell population marked with a $CD44^{hi}CD24^{low}$ phenotype, which is similar to that observed in epithelial stem cells[109]. Moreover, carcinomas contain a subpopulation of self-renewing, tumour-initiating cells, known as cancer stem cells (CSCs), which efficiently generate new tumours. In mammary carcinomas, induction of EMT promotes the generation of $CD44^{hi}CD24^{low}$ CSCs that are able to form mammospheres, and similarly-defined CSCs isolated from tumours express EMT markers[109]. Consistent with the reversible nature of EMT, differentiated cancer cells can transition into CSCs, and vice versa, enabling oncogenic mutations that arose in

differentiated cells to integrate through EMT into CSCs. As EMT promotes cell invasion that leads to tumour cell dissemination, CSCs with new oncogenic mutations could clonally expand, following invasion, dissemination and MET in secondary tumours[100,110].

LOXL2 has been related to EMT in various ways. For example, EMT inductors (such as hypoxia and TGFβ) promote LOXL2 expression[111,112]. LOXL2 contributes to the EMT process by repressing E-cadherin together with SNAIL1. Moreover, LOXL2 can interact and stabilize SNAIL1 (a master regulator of EMT) and even induce EMT by itself[90]. Since LOXL2 has been shown to be key in EMT process, elucidating the role of the LOXL2 protein in all EMT related functions, including embryonic development regulation, stem cell properties and cancer progression, will be critical to understanding how these processes are regulated.

2.4. LOXL2 and cell differentiation

Some studies have linked LOXL2 expression to regulation of cell differentiation. For instance, during keratinocyte differentiation, LOX expression is increased, while LOXL2 transcription is downregulated[113]. This work suggests a switch in the expression of both proteins: LOXL2 would be expressed in progenitor cells but, upon keratinization, it would

be repressed and LOX upregulated. In fact, silencing of LOX expression inhibits differentiation, in contrast to LOXL2 silencing.

Concomitantly, the progression of squamous cell carcinoma of the skin is associated with enhanced expression of LOXL2. These tumour cells derive from keratinocyte stem cell precursors and are characterized by aberrant differentiation, as they have express reduced levels of differentiation markers[114]. In this case, LOXL2 overexpression observed in squamous cell carcinomas may contribute to tumour progression through inhibition of the differentiation of keratinocyte-derived tumour cells.

Moreover, in adult human dental pulp stem cells (hDPSCs) undergoing differentiation to odontoblast-like cells, LOXL2 levels are reduced at both the mRNA and protein levels. However, the levels of other LOX family members, including LOX, LOXL1, LOXL3 and LOXL4, are increased. Indeed, LOXL2 has a negative effect on the differentiation of hDPSCs, since blocking its activity promotes hDPSC differentiation to odontoblasts[115].

In contrast, when studying a model of chondrocyte differentiation (ATDC5 cells), LOXL2 expression increases significantly, while LOX expression decreases. Moreover, when LOXL2 expression was depleted with shRNA

constructs, chondrocyte differentiation was abolished, pointing to a key role for LOXL2 in this process[116]. These studies were performed in adult stem cells, which are multipotent stem cells. Multipotent cells are less plastic and more differentiated stem cells that give rise to a limited range of cells within a tissue type and represent the offspring of the pluripotent cells that become progenitors of different cell lines, such as intestinal stem cells, skin epithelial stem cells or hematopoietic stem cells[117]. Hence, LOXL2 function, in contrast to that of the other LOX family members, is related to progenitor cell features, since it is expressed in adult stem cells in different tissues and its expression is lost upon final differentiation.

3. Embryonic stem cells (ESCs)

During embryonic development, a single cell, the zygote, is able to give rise to all cell types for the formation of a new organism (*Figure I.13*). This cell undergoes a serial of divisions after fertilization to generate a homogenous cell mass, the morula. Subsequently, these cells continue to divide to the blastocyst stage, to generate two cell populations with distinct fates and developmental potential: the inner cell mass (ICM), which will contribute to the formation of the embryo; and the trophectoderm (TE), which gives rise to extraembryonic tissues supporting placenta development[118]. Afterwards, ICM cells further split into two groups: the epiblast, a pluripotent embryonic layer; and the primitive endoderm (PE), which contributes to extraembryonic tissues.

At the onset of gastrulation, groups of cells from the epiblast move to the primitive streak (PS) and undergo EMT to form the mesendoderm, a transient population that is then separated into endoderm and mesoderm[119]. The remaining epiblast cells give rise to the embryonic neuroectoderm, thus establishing the three primary germ layers, which will develop into all mature cell types in the animal.

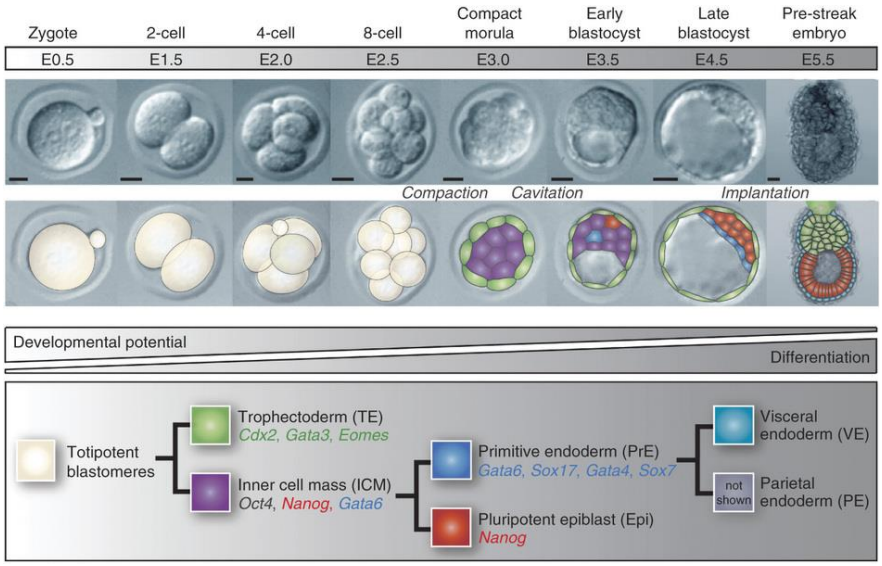

Figure I.13. Overview of early embryonic development. Lineage segregation before implantation implicates two cell fate decisions. The first one gives rise to trophectoderm and inner cell mass, and the second one leads to the allocation of primitive endoderm and epiblast. Gene expression pattern of each lineage is noted below each cell type. Post-implantation, primitive endoderm differentiates into visceral and parietal endoderm. E stands for embryonic day. Scale bars indicate 50 □m[120].

Mouse ESCs were obtained for the first time from explanting blastocysts or ICMs on a feeder layer of inactivated fibroblasts[121]. Stem cells are defined by their capacity for repeated generation of two classes of progeny, daughter cells with equivalent proliferative and developmental potential and daughters specified for differentiation. Since ESCs have an indefinite proliferative life span, and long-lived subclones obtained by single-cell expansion retain pluripotency, they can be described as self-renewing stem cells.

Classical culture conditions employed serum-containing media and a layer of mitotically-inactivated fibroblasts (feeder) cells[122]. Initially, little was known about the molecular nature of the self-renewal signals provided by these components. However, further studies discovered that contribution of feeders was determined to be the IL-6 family cytokine leukaemia inhibitory factor (LIF)[123,124], creating a feeder-free culture condition for mESCs.

The apparent dependence of ESCs on growth factors pointed to a control of ESC maintenance by exogenous signals and downstream activated signalling pathways. Following this idea, it was found that inhibiting glycogen synthase kinase 3 beta (Gsk3β) and fibroblast growth factor (FGF)-MAPK pathways with two different inhibitors (2i) permits long-term propagation of mESCs, exhibiting also enhanced cloning efficiency when adding LIF to the media[125]. This optimized condition, 2i+LIF, has since been used to derive ESCs from previously recalcitrant strains and species.

ESCs can also differentiate to form cells of the three germ layers *in vitro* and contributes to all tissues in chimeric mice generated by blastocyst injection[126], meaning that they are pluripotent. The molecular events and timeline of ESC differentiation *in vitro* faithfully recapitulate embryonic development *in vivo*. Hence, ESCs represent a powerful

model system to study pluripotent state and developmental cell differentiation[127].

3.1. Transcription network maintaining pluripotency

The POU-domain transcription factor Oct4 (Pou5f1, also known as Oct3) is the preeminent pluripotency factor and was the first transcription factor identified and characterized as a regulator of pluripotency[128]. It is essential for ICM identity, since morulas lacking this protein have inner cells that differentiate along TE lineage and do not develop any embryonic rudiment[129]. In established ESCs, deletion of Oct4 leads to loss of self-renewal and differentiation. However, forced expression of Oct4 does not consolidate or enhance ESC self-renewal. In fact, even modest overexpression precipitates differentiation[130], suggesting a dual role for Oct4 in self-renewal and differentiation depending on its expression levels, which must be regulated tightly.

The SRY-box transcription factor Sox2 is also essential for ESC self-renewal. Sox2 acts together with Oct4 to regulate multiple target genes, including key transcription factors such as Nanog. Sox2 inactivation in ESCs results in trophoblast formation, phenocopying Oct4 deletion[131]. Both proteins interact and bind to DNA at Oct/Sox elements[132], positively regulating Oct4 expression. In addition to the epiblast, Sox2

is also expressed in the trophectoderm and, later on, in neuroectodermal cells[133]. Similar to Oct4, Sox2 overexpression predisposes ESCs to differentiation[134].

The Nanog transcription factor is also a classic pluripotency-maintaining factor. It is not homogenously expressed throughout the ICM but rather is restricted to the epiblast at the late blastocyst stage[135]. Embryos with depleted Nanog fail to establish an epiblast due to ICM degeneration[136]. However, ESCs lacking Nanog can self-renew and retain pluripotency in optimal culture environment containing 2i+LIF, but have an impaired colony-forming capacity in just LIF-containing media[137]. When overexpressed, Nanog confers the ability to self-renew to ESCs in the absence of LIF[138].

Evidence suggests that Oct4/Sox2 and Nanog have complementary and only partially overlapping gene regulatory activities, since gene expression profiling after knockdown of either Oct4 or Nanog shows distinct transcriptional responses[139]. In addition to Oct4, Sox2 and Nanog, other transcription factors also reinforce ESC identity, including the well-studied Klf4, Esrrb, Klf2 and Tbx3. These transcription factors are specific to naïve epiblasts and ground-state ESCs and are downregulated in post-implantational epiblasts and at the onset of ESCs differentiation[140]. Moreover, they are interconnected by regulatory loops with one another as well as with Oct4, Sox2 and Nanog (*Figure I.14*).

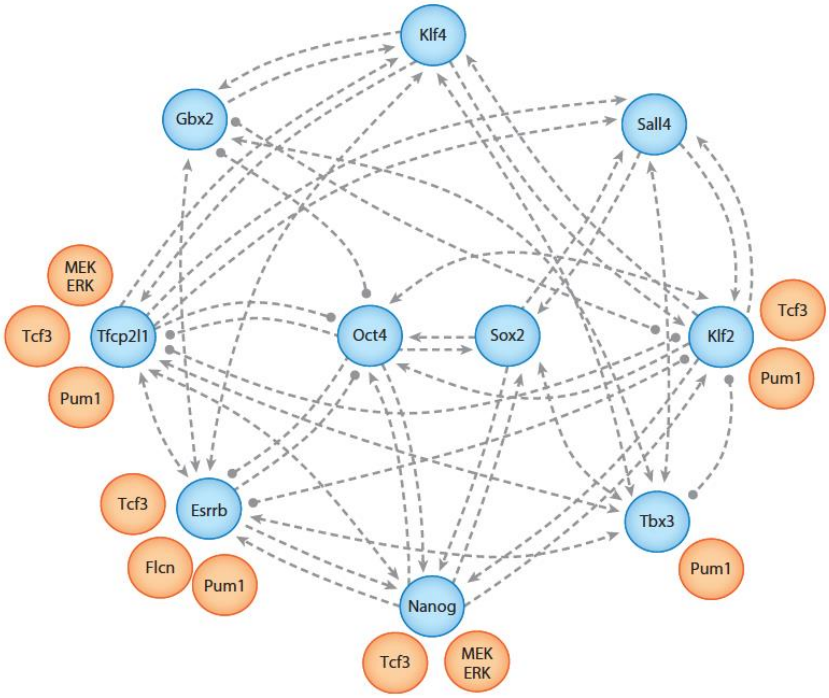

Figure I.14. Pluripotency transcriptional network. Nodes of the ground state pluripotency circuit are indicated in blue. These factors can be targeted by destabilizers (negative regulators), shown in orange. Dashed lines indicate potential interactions between nodes inferred from correlated expression[141].

This transcription network is sufficient to reprogram mature differentiated cells back to an ESC-like state when reinitiated through overexpression of only four factors: Oct4, Sox2, Klf4 and c-Myc, also known as the Yamanaka factors[142]. The resulting iPS cells are transcriptionally, epigenetically, and functionally equivalent to ESCs. Addition of other factors to the cocktail, such as Nanog, Tfcp2l1, and Sall4, can increase reprogramming efficiency[143]. Therefore, the factors key to propagation of the pluripotent ground state can also contribute to regenerating this state *in vitro*, and when

pluripotent cells commit to differentiate towards specific lineages, this self-sustained transcription network is terminated.

3.2. Lineage commitment of the mesoderm and endoderm

The formation of primitive streak (PS) in the epiblast is the first stage of lineage commitment. After embryo gastrulation, primary germ layers are generated and basic body plan in the embryo is established[144]. Groups of cells move from epiblast through the PS where they undergo EMT and emerge as different mesoderm or endoderm tissues depending on their location regarding the anterior-posterior axis[145]. Dissolution of the naïve pluripotency program and specification of germ cells and gastrulation proceeds in less than 24 h, suggesting a deterministic molecular program.

However, ESCs have been proposed to start differentiation stochastically due to fluctuations in expression of several transcription factors such as Nanog[146]. In fact, ESCs in culture with LIF show mosaic expression of Nanog and other pluripotency transcription factors[137]. Expression seems to fluctuate in a dynamic equilibrium[147]. Various gene expression patterns have been proposed to explain how ESCs sample different cell fate options, which could be

essential for pluripotency maintenance. Although this is not observed in ESCs cultured with 2i, these are also pluripotent and differentiate properly both *in vitro* and *in vivo*. It is likely that the ESC heterogeneity in the presence of LIF is due to an epiphenomenon, a mixture of transitory cell states generated in response to incoherent environmental inputs[148].

ESCs undergo unidirectional developmental progression upon 2i withdrawal, phenocopying the behaviour of peri-implantation epiblast. However, this is not instantaneous, as all cells conserve self-renewal capacity for 24 hours[149]. Hence, an exit from the ground state is reversible until cells reach a transition point, at which ESC identity is permanently extinguished.

Some pluripotency factors, such as Klf5 and Sox2, are also expressed in specific lineages, suggesting that pluripotency could be a precarious balance in which rival lineage specifiers constantly compete[150]. This competition seems to remain in balance during the ground state and to be resolved during differentiation, when one family of factors then dominates over the others (*Figure I.15*).

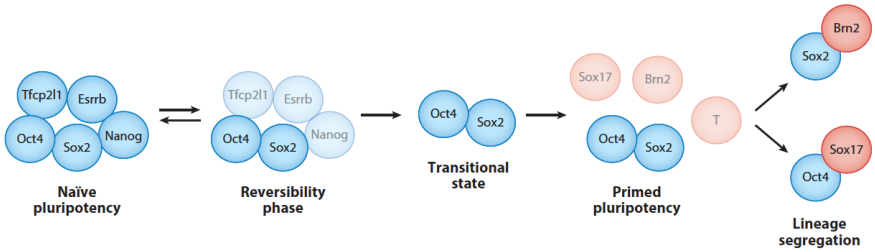

Figure I.15. Developmental progression from ESCs to lineage commitment. During progression to lineage commitment, ESCs are released from ground state maintenance. Initial phase factors are represented in blue. Lineage commitment factors, in orange, are Sox2 and Brn2 for neuroectodermal progenitors, and Oct4 and Sox17 for mesendodermal progenitors[141].

ESCs do not enter into lineage commitment directly as they exit self-renewal; rather, they first must lose expression of all ground state factors. During the transition, they upregulate the expression of certain transcription factors, such as Otx2, which are increased immediately after implantation *in vivo*, but they do not express lineage-specific factors yet. This transition state seems to represent the early post-implantation epiblast stage[151].

Oct4 is the only pluripotency factor known to be continuously expressed throughout the pre- and post-implantation epiblast. Oct4 interacts with many other transcription factors, which suggests that Oct4 may serve as a scaffold for multiple gene expression programs that initiate alternative lineage commitment.

3.3. Neural lineage commitment

Epiblast cells that are not recruited to the PS will form the neuroectoderm, with a commitment that is completed at the end of gastrulation[152]. In neural induction, a nervous system is generated from neuroectodermal precursors on the dorsal side of the embryo in response to signals from adjacent dorsal mesoderm. By the end of gastrulation, the neuroectoderm cells form the neural plate, which subsequently folds into a tube before developing into the brain at its interior end, and the spinal cord at its posterior end.

ESCs have two principle approaches to induce neural identity. The first one is based on inhibiting mesendoderm-inducing SMAD signalling pathways (activated by Nodal and Bmp signals) and the activity of endogenous FGF signals[153]. This approach resembles the signalling pathways that occur during embryo development. Inhibition of Nodal pathway downregulates Nanog and promotes Zeb2 expression, which directly represses Oct4 expression and promotes expression of other neural-specific genes[154]. Concomitantly, Bmp inhibition supresses the induction of non-neural germ layers, maintaining the expression of neural genes such as Sox2[155]. Subsequently, FGF and Bmp/Nodal inhibitors consolidate neural lineage commitment by downregulating the pluripotency transcription network in ESCs, preventing the

induction of mesendoderm determinants and promoting the onset of neural transcription programs.

The second approached is based on treatment of ESCs cultured as suspension aggregates, termed embryonic bodies (EBs), with retinoic acid (RA). The changes in morphology and gene expression that EBs undergo after several days of continued culture resembles *in vivo* neural development. After RA induction, expression of the Oct4 pluripotency gene is rapidly downregulated, and neuroectodermal markers such as Sox1 start to be upregulated. Subsequently, EBs start to express neural progenitor markers such as Pax6 and Nestin. Finally, post-mitotic neurons expressing Tuj1 and NF appear in the culture (*Figure I.16*). Similar to the first method, neural lineage commitment in this case also involves downregulating the pluripotency transcription network, initiating neural transcription programs and suppressing alternative fates.

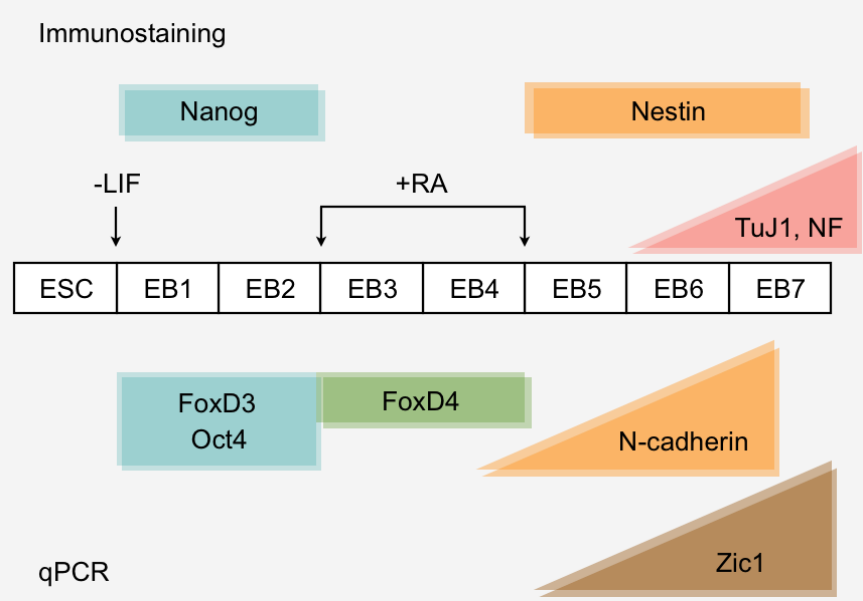

Figure I.16. Gene expression pattern in neuronal differentiation. When EBs are treated with RA for two days and then cultured for an additional 3–5 days without RA, they differentiate into neurons expressing a class III β-tubulin (TuJ1) marker. Markers detected by immunostaining are represented in the protocol scheme. Gene expression results from qPCR assays are shown below[156].

The ability of RA to induce neural differentiation can be harnessed to produce specific neural cell types that can then be used for therapeutic transplantation. ESCs, haematopoietic stem cells and neural stem cells can be diverted down the neural differentiation route using combinations of RA and growth factors or neurotrophins (*Table I.1*).

Cell type	Inducers	Neuronal type
Human/mouse ESCs	RA + SHH	cholinergic, dopaminergic
Mouse ESCs	RA + CNTF	dopaminergic
Human ESCs	RA + BDNF, RA + TGFα	dopaminergic
Mouse embryonic stem cells	RA	Glutaminergic
Adult neural stem cells	RA + NT-3	Mixed
	RA + KCl	GABAergic
Human olfactory neural cells	RA + SHH	Dopaminergic
Bone marrow haematopoietic cells	RA, RA + NT-3, RA + BDNF, RA + FGF	NS
	RA + SHH	Glutaminergic

Table I.1. Neural types induced by RA with or without other stem cell factors. BDNF, brain-derived neurotrophic factor; CNTF, ciliary neurotrophic factor; FGF, fibroblast growth factor; KCl, potassium chloride; NS, not specified; NT-3, neurotrophin-3; RA, retinoic acid; SHH, sonic hedgehog; TGFα, transforming growth factor α[157].

Some of these combinations have been tested *in vivo* for their ability to replace lost neurons. For example, embryonic neural progenitor cells differentiated with RA have survived and become neurons when grafted into a range of locations in adult brain[158]. The potential of such differentiated cells might thus be remarkable.

Most available protocols for neural differentiation from ESCs are based on EB formation (as explained above) or on co-culturing ESCs with particular stromal cell lines, including PA6 and MS5. However, ESCs under an adherent monoculture system not involving EB formation or co-culture are

successfully driven into a neural fate when treated with RA[159]. Exposure of ESCs to RA in the initial 2 days of differentiation leads to efficient generation of neural progenitors and further differentiation to neuronal fate.

3.3.1. RA signalling pathway

RA is a metabolic product of liposoluble vitamin A (retinol). As most animals cannot synthesize vitamin A, they obtain it from their diet and store it as retinoids in the liver[160]. Retinoids are transported as retinol, which is released into the bloodstream and bound to retinol-binding protein 4 (RBP4) before taken up by target cells via membrane receptor for RBP4, STRA6[161]. Once in the cytoplasm, retinol bind to retinol-binding protein 1 (RBP1) and is converted to *all-trans* RA by retinol dehydrogenase and retinaldehyde dehydrogenases (RALDHs)[162]. The newly-synthesized RA is then bound by cellular retinoic acid binding proteins 1 and 2 (CRABP1 and CRABP2) in the cytoplasm (*Figure I.17*).

RA can then act in a paracrine manner, being released from the secreting cells and taken up by receiving cells, or it can act in an autocrine fashion. Either way, RA enters the nucleus with the assistance of CRABP2 and binds to heterodimers of ligand-inducible transcription factors comprising the RA

receptors (RARα, RARβ and RARγ) and retinoic X receptors (RXRα, RXRβ and RXRγ)[163].

RA-bound RAR/RXR complexes regulate gene expression by binding to DNA sequences within the promoter of target genes called retinoic acid response elements (RAREs)[164]. After RA has activated RARs, it exits the nucleus and is catabolized by the CYP26 class of P450 enzymes in the cytoplasm[165].

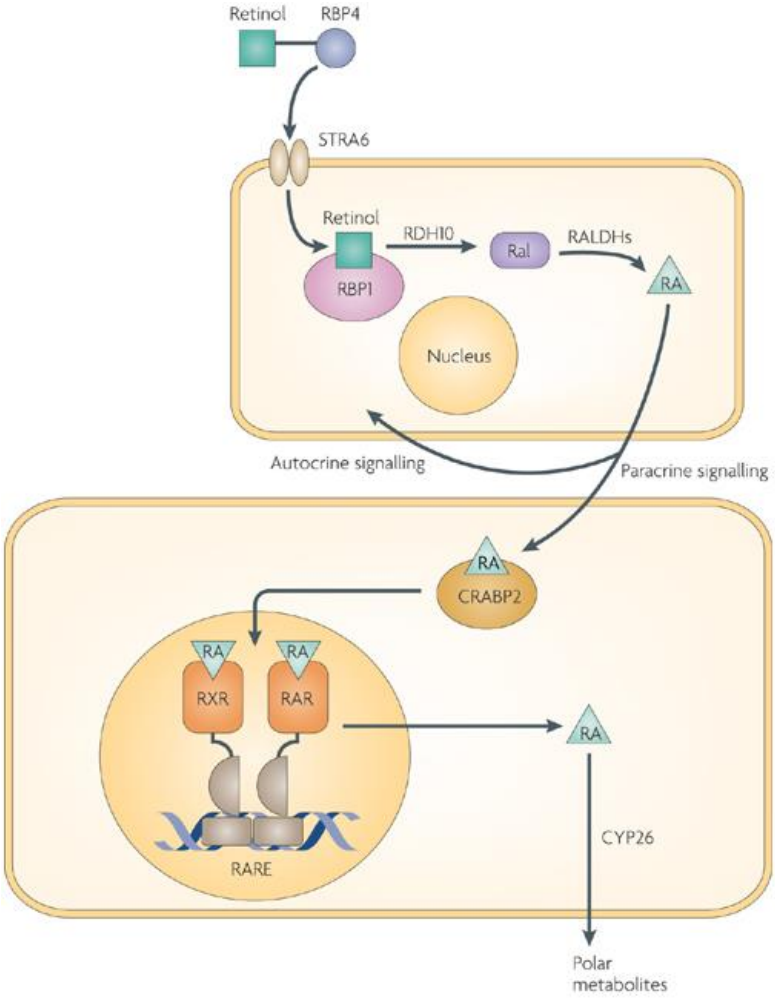

Figure I.17. Pathways involved in the generation, action and catabolism of retinoic acid (RA). Retinol bound to RBP4 is taken up by cells through STRA6 receptor. RDH10 metabolizes retinol to retinaldehyde (Ral), which is then metabolized to RA by RALDHs. RA can be released from the cytoplasm and taken up by the receiving cell (paracrine signalling) or can act back on its own nucleus (autocrine signalling). CRABP2 assists RA entry into the nucleus. In the nucleus, RA binds to RA receptors (RARs) and retinoid X receptors (RXRs), which heterodimerize and bind to DNA sequence known as the retinoic acid-response element (RARE). This binding activates transcription of target genes. RA is then catabolized in the cytoplasm by the CYP26 class of P450 enzymes[157].

It has been proposed that in the absence of RA, the apo-receptor heterodimer RAR/RXR binds to the RAREs and recruits corepressors and histone deacetylase complexes (HDACs) to maintain target gene repression[166]. In the presence of ligand, a conformational change leads to the replacement of corepressors by coactivator complexes (*Figure I.18*). This induces chromatin remodelling and facilitates the assembly of the transcription pre-initiation complex, therefore activating target gene expression[167].

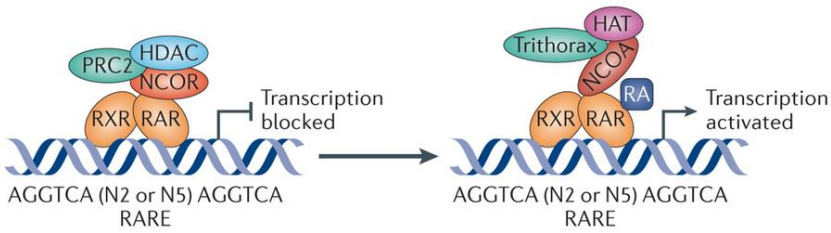

Figure I.18. RA signalling mechanism related control of transcription. The absence of RA allows co-repressors of the nuclear receptor corepressor (NCOR) family to bind RAR and recruit repressive factors such as Polycomb repressive complex 2 (PRC2) and histone deacetylase (HDAC), whereas the presence of RA releases co-repressors and allows co-activators of the nuclear receptor co-activator (NCOA) family to bind to RAR and recruit activating factors such as histone acetylases (HAT)[168].

3.3.1.1. RA regulated genes in ESC differentiation

RA treatment of ESCs results in critical changes in gene expression regarding not only for transcription factors but also for metabolism-related proteins, extracellular matrix components, proto-oncogenes, growth factors, cytoskeletal proteins, cell cycle regulators and apoptosis related factors, among others[169].

Genes are classified into two different categories: early responding and late responding genes. Early responding ones are induced within the first 8–16 hours and usually have binding sites for RA receptors in their promoters; these include several Hox genes, among others[170]. Late response genes can be divided into two additional categories. The first group, which is larger, contains genes that change their expression in one or more days. Rex1, NeuroD and N-cadherin, and others noteworthy genes related to specific differentiation pathways are in this group and appear to be regulated by RA indirectly. The second group contains genes marking terminal differentiation and are expressed 5–6 days after RA stimulation.

3.3.1.2. RA repression of pluripotency transcription network

The first step in ESC differentiation is terminating the transcription network that sustains the pluripotent state. Upon RA treatment, key pluripotency factors, such as Oct4 and Nanog, are rapidly downregulated at both mRNA and protein levels[171]. Since RA receptors act almost exclusively as transcriptional activators, repression of pluripotency genes is likely to be an indirect effect of RA mediated by transcriptional repressors downstream of RA pathway. According to this idea, Oct4 mRNA levels undergo little changes in the first 8 hours after RA treatment, with most downregulation occurring 8–24 hours post-RA exposure[170]. Moreover, the Oct4 promoter does not have RA receptor binding sites at its promoter, reinforcing the idea that RA regulation is indirect.

One of the first RA-related elements to be identified as regulator of pluripotency factor Oct4 are the several hormone response elements (HREs) that are found in Oct4 enhancer regions. Nuclear orphan receptors Coup-TFs, which are upregulated upon RA treatment, bind these elements and mediate repression[172]. Another transcription factor related to Oct4 downregulation is the RA primary response gene Hoxa1, which has a RA receptor binding site in its 3′ enhancer. RA treatment of Hoxa1 null ESCs results in higher Oct4 mRNA levels as compared to wild-type cells. Moreover,

when these cells were forced to differentiate upon LIF withdrawal, they failed to differentiate into neurons and followed an endodermal differentiation pathway[173].

Lastly, the orphan nuclear receptor GCNF has also been proposed to repress pluripotency genes upon RA differentiation[174]. GCNF is a transcription repressor that binds directly to promoters of target genes. During development, GCNF expression pattern is inversely correlated with that of Oct4[175]. Interestingly, GCNF protein levels are induced by RA treatment after 24–36 hours and drop to undetectable levels after 3 days, while Oct4 protein levels decrease between 24–72 hours after induction. GCNF can bind response elements located in Oct4 and Nanog promoters. Concomitantly, among other pluripotency factors, Oct4, Nanog and Sox2 expression is not efficiently turned off in GCNF mutant ESCs.

To fully understand the mechanisms by which RA regulates pluripotency transcription network in ESCs, it is crucial to identify which of the early response genes are involved in the repression of the pluripotency genes.

OBJECTIVES

Lysyl oxidase-like 2 (LOXL2) has been described to have different roles in several cellular functions. However, the molecular mechanisms by which LOXL2 acts in such different aspects of cell function are still unknown.

The general objective of this thesis is thus to describe new roles for LOXL2 enzyme. To this aim we focused on:

I. Screening of new LOXL2 substrates
II. Characterization of most significative candidates
III. Finding biological relevance of oxidation of these substrates by LOXL2.

RESULTS

1. Characterization of new LOXL2 substrates

Many posttranslational modifications performed by histone-modifying enzymes occur not only in histones but also in non-histone proteins, which can regulate protein function by regulating protein-protein interactions, protein stability, localization, and enzymatical activities[176]. We therefore asked whether LOXL2 might also have non-histone substrates, and if so, what are those proteins and how did oxidation affect their function.

To find new putative substrates, we used an unbiased proteomic approach consisting of a biotin-hydrazide pull-down protein purification. Biotin-hydrazide is an activated biotin that reacts with aldehyde groups[91]; hence, sample incubation with this compound enables tagging of proteins that present an aldehyde group. Since the LOXL2 reaction leaves an aldehyde in the product lysine residue, possible substrates of this enzyme should also be tagged with a biotin in the presence of biotin-hydrazyde.

First, nuclear extracts from HEK293T cells, transfected with wild-type (LOXL2 wt) or an inactive (LOXL2 mut) form of LOXL2, were incubated with biotin-hydrazide to label oxidized proteins. We had previously generated in the laboratory a double-point mutant for LOXL2 in which two histidine residues of the catalytic domain involved in copper binding

were mutated to glutamine (H626Q, H628Q)[91]. However, as this mutant still presented residual catalytic activity, we also mutated the tyrosine residue involved in the formation of lysyl tyrosylquinone cofactor to phenylalanine. The resulting LOXL2 mutant thus has the H626Q, H628Q and Y689F mutations (*Figure R.1a*).

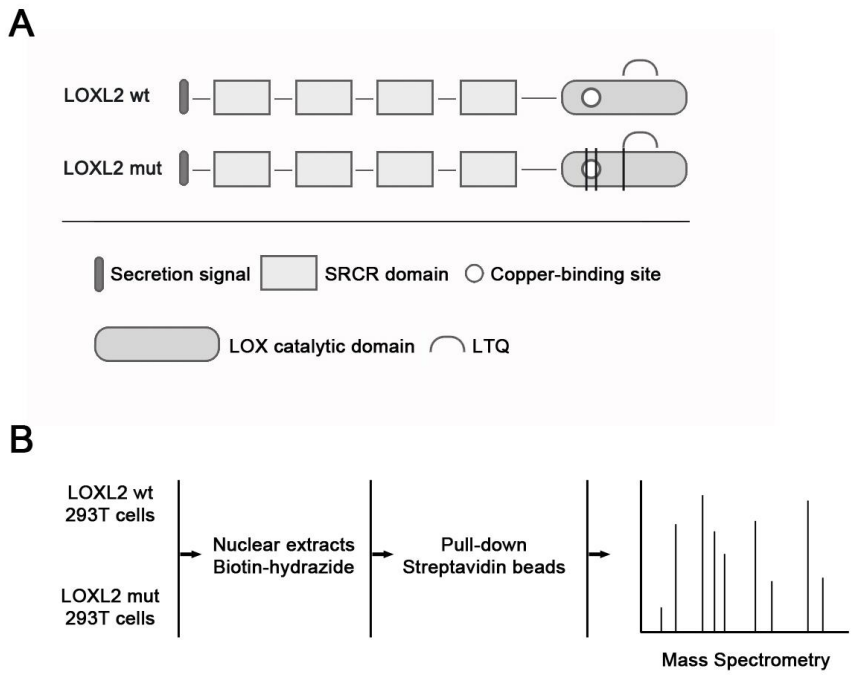

Figure R.1. Experimental approach to identify LOXL2 substrates. (**A**) Cartoon representing LOXL2 protein showing its main domains and point mutations in the mutant form. (**B**) Scheme of the experimental procedure used to identify new substrates by mass spectrometry.

Samples were then incubated with streptavidin beads to pull down oxidized proteins, and the precipitated fraction was eluted with urea. To identify proteins, samples were digested with trypsin and analysed by high-resolution mass spectrometry (MS) (*Figure R.1b*).

We identified 117 candidate proteins that were specifically enriched in cells transfected with wild-type LOXL2 as compared to cells expressing the catalytically inactive mutant (Table A.1). Among these putative substrates, we identified histone H3, a previously described substrate for LOXL2[91]. Interestingly, when performing a STRING analysis of the putative substrates, we found the TFIID transcription factor complex as an enriched network (Figure R.2). Search in the KEGG pathway database to identify the main represented biological processes also revealed a significant enrichment for basal transcription factors (Table A.2).

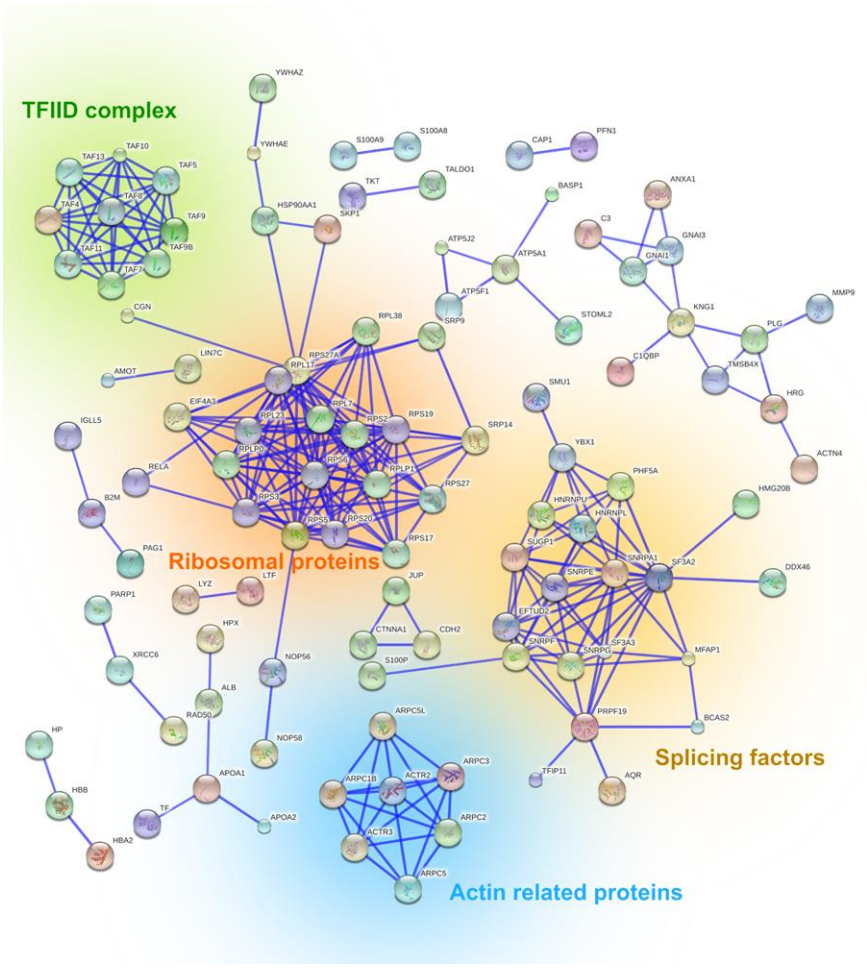

Figure R.2. STRING analysis of new LOXL2 putative substrates. Representation of interactome analysis performed for the new putative substrates by STRING. Only the highest confidence interactions are shown (0.900).

TAF10 protein, one of the TFIID complex subunits identified, was already characterized to have a methylated lysine[177]. Since LOXL2 has a strong preference for methylated histone H3[91], we thus focused our studies on TAF10 as a possible LOXL2 substrate. We reasoned that LOXL2 could have a dual role in transcriptional repression: on the one hand, it

could deaminate histone H3K4me3, as initially described[91], and on the other hand, it could oxidize methylated TAF10, which could impair the transcription of TFIID (methylated TAF10)–dependent genes[177].

Interestingly, we found that the histone methyltransferase responsible for TAF10 methylation, SET7/9, was also described to methylate histone H3 in lysine 4[177,178]. Since both SET7/9 and LOXL2 could share two different substrates, we reasoned that they could also share a recognition motif in substrate proteins. Indeed, a recognition motif for SET7/9 methyltransferase had already been proposed[179]; thus, we looked for this motif in the 117 putative LOXL2 substrates. Surprisingly, 20 proteins, representing 17% of the total candidates, showed one or more putative SET7/9 recognition motifs (*Table A.3*), suggesting that these two enzymes may indeed share recognition motif and target residues.

To validate the MS assay, we first confirmed that TAF10 was indeed oxidized and enriched in nuclear extracts when the wild-type form of LOXL2 was expressed as compared to the catalytically inactive LOXL2 (*Figure R.3a*). Moreover, co-immunoprecipitation experiments showed that ectopically expressed LOXL2-FLAG and TAF10-HA proteins interact in HEK293T cells (*Figure R.3b*).

Interestingly, we observed three bands interacting with LOXL2 (*Figure R.3b, middle panel*), suggesting LOXL2 interacts with a posttranslationally modified form of TAF10. We confirmed this possibility by checking that those three bands corresponded to TAF10 using a specific monoclonal antibody for TAF10 protein (*Figure R.3b, lower panel*).

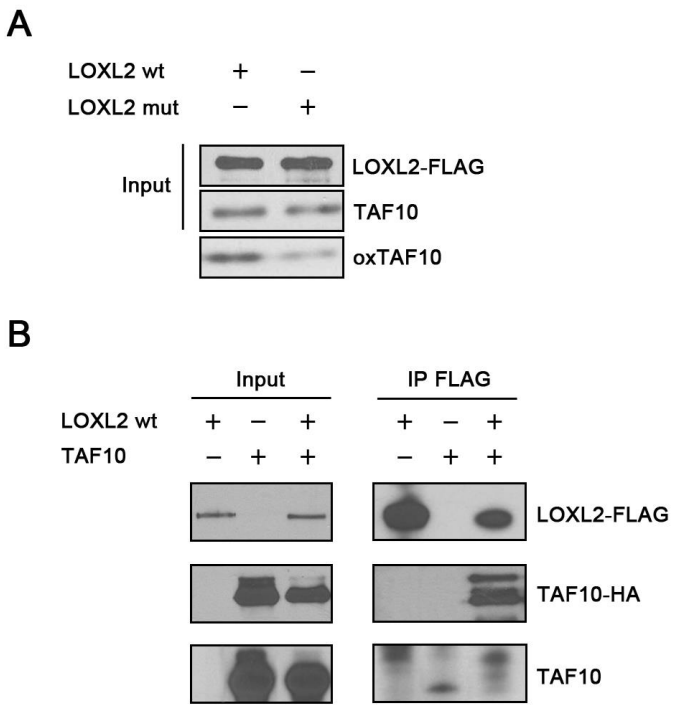

Figure R.3. LOXL2 interacts with TAF10 and increases its oxidation levels. (**A**) Oxidized proteins from HEK293T cells transfected with LOXL2 wt-FLAG or LOXL2 mut-FLAG were pulled down with streptavidin beads. The oxidation levels of TAF10 were determined by Western blot with an anti-TAF10 antibody. (**B**) Proteins were co-immunoprecipitated from HEK293T cells transfected with LOXL2-FLAG or/and TAF10-HA with anti-FLAG antibodies. Precipitates were analysed by Western blot for LOXL2-FLAG, TAF10-HA and TAF10.

2. LOXL2 interacts with a highly posttranslationally modified TAF10

To characterize which modifications TAF10 presents when interacting with LOXL2, we analysed the co-immunoprecipitated TAF10 fraction by MS and looked for posttranslational modifications. As expected, TAF10 was posttranslationally modified (*Table A.4*). Strikingly, most modifications detected in TAF10 were localized in the histone-fold domain of the protein (*Figure R.4a*), which is crucial for the interactions of TAF10 with other subunits of the TFIID complex[180]. We identified several putative posttranslational modifications for TAF10, including serine phosphorylation, lysine monomethylation, lysine dimethylation, lysine trimethylation and lysine oxidation (*Figure R.4a*). Among all the observed modifications, only the lysine 189 (K189) methylation had been previously described; importantly, it was also reported that mutations in this residue abolish methylation of TAF10[177].

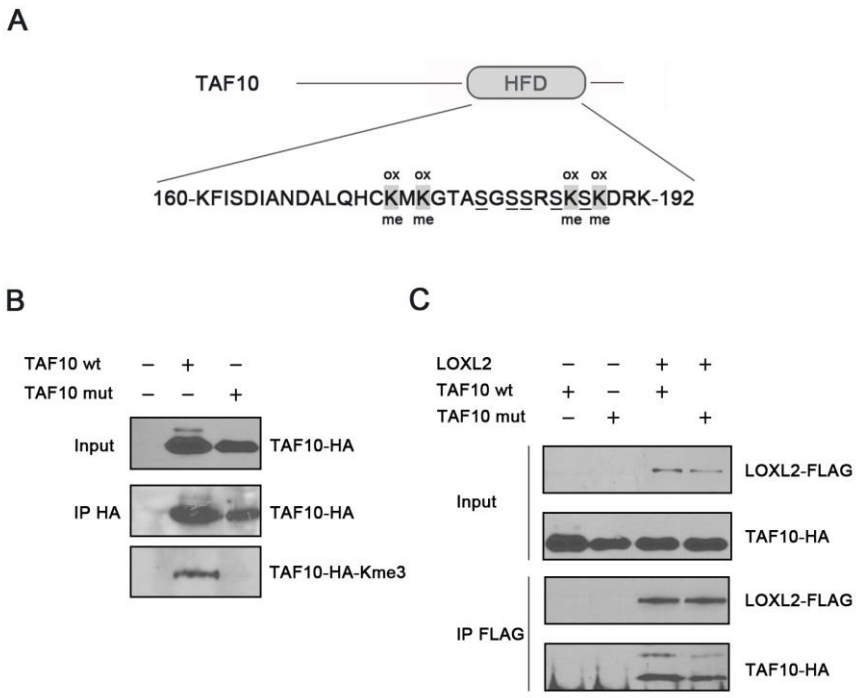

Figure R.4. LOXL2 interacts with a highly posttranslationally modified TAF10. (**A**) Representation of the PTMs identified in ectopic TAF10 coimmunoprecipitating with LOXL2. Phosphorylated serines are underlined, while methylated and oxidized lysine residues are highlighted in grey. (**B**) HEK293T cells transfected with pcDNA3, TAF10-HA and/or TAF10K189Q-HA were analysed by protein immunoprecipitation with anti-HA antibodies and then re-immunoprecipitated with an anti-trimethyl lysine antibody. Levels of trimethylated TAF10 were analysed by Western blot against HA. (**C**) HEK293T cells transfected with LOXL2-FLAG, TAF10-HA and/or TAF10K189Q-HA were subjected to protein coimmunoprecipitation analyses with anti-FLAG antibodies and Western blotted for LOXL2-FLAG or TAF10-HA.

We further confirmed that TAF10 was trimethylated by immunoprecipitation and Western blot using an antibody that recognizes trimethylated lysines. Ectopically expressed TAF10-HA, either wild-type (TAF10 wt) or K189 mutated to glutamine (mut), was immunoprecipitated with antibodies

against HA tag, re-immunoprecipitated with antibodies against trimethylated lysines, and analysed by Western blot with antibodies against HA. Trimethylated TAF10 (TAF10-HA-Kme3) was detected only in cells transfected with TAF10 wt (*Figure R.4b*), verifying that the methylation of TAF10 depends on K189 as previously described[177]. Interestingly, depletion of methylation in TAF10 did not affect its interaction with the LOXL2 protein, since both TAF10 wt and TAF10 mut proteins coimmunoprecipitated with LOXL2 when ectopically expressed (*Figure R.4c*).

3. LOXL2 oxidizes methylated TAF10

To validate that methylated TAF10 was oxidized by LOXL2, we co-transfected wild-type LOXL2 with either wild-type TAF10-HA or the methylation-defective K189Q mutant. Oxidized proteins were immunoprecipitated and analysed by Western blot against an HA tag (*Figure R.5a*). We observed that LOXL2 oxidized the methylated wild-type TAF10 as well as, to a much lesser extent, the nonmethylated TAF10 mutant. Concomitantly, the global methylation levels of TAF10 decreased in the presence of wild-type LOXL2 but not when catalytically inactive LOXL2 was transfected (*Figure R.5b*).

We also analysed LOXL2 activity on methylated TAF10 *in vitro* by incubating recombinant LOXL2 protein, produced in baculovirus, with immunoprecipitated wild-type—and hence methylated—TAF10. Mutant LOXL2 was heat denatured to eliminate any residual activity and used as a control of the reaction. After the reaction, biotin-hydrazide was added to the product, and oxidized TAF10 was re-immunoprecipitated with streptavidin beads and analysed by Western blot against an HA tag. TAF10 methylation levels were also checked after the reaction by re-immunoprecipitating the product with an anti-trimethyl lysine antibody. We observed that recombinant LOXL2 directly oxidized modified TAF10 *in vitro* with a

concomitant decrease in trimethylation levels of TAF10 (*Figure R.5c*).

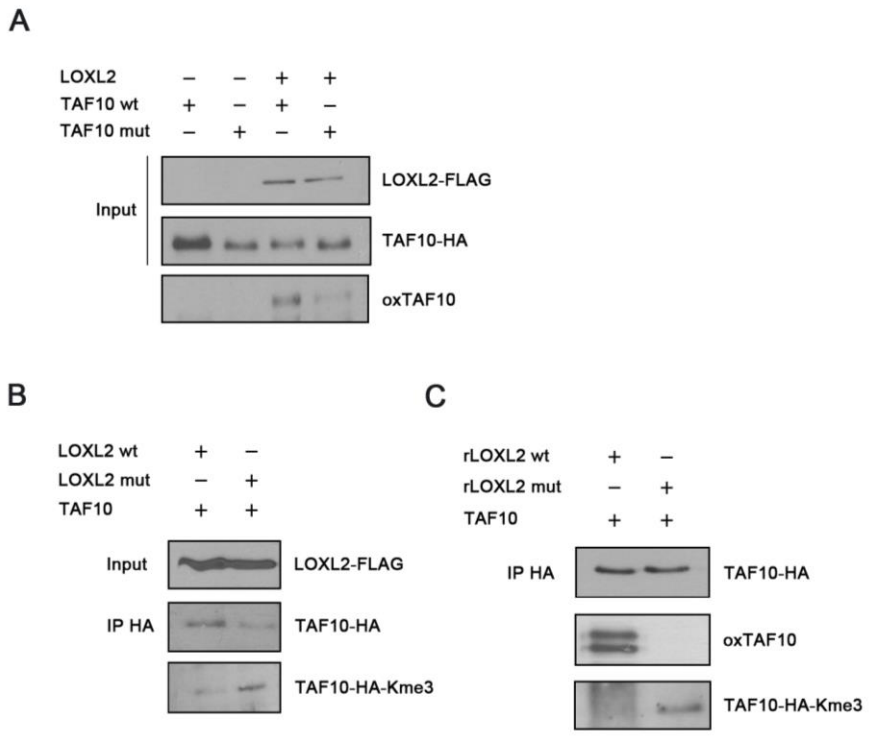

Figure R.5. LOXL2 oxidizes TAF10 and reduces its trimethylation levels. (**A**) Cell lysates from HEK293T cells transfected with LOXL2-FLAG or/and TAF10-HA or TAF10K189Q-HA were incubated with biotin hydrazide, and biotin-labelled proteins were pulled down with streptavidin beads. The oxidation levels of TAF10 were determined by Western blot with an anti-HA antibody (oxTAF10). (**B**) Cell lysates from HEK293T cells transfected with LOXL2 wt-FLAG or LOXL2 mut-FLAG and TAF10-HA. Methylated levels of TAF10 were analysed by protein immunoprecipitation with anti-HA antibodies and then re-immunoprecipitated with an anti-trimethyl lysine antibody. The levels of trimethylated TAF0 were analysed by Western blot against HA (TAF10-HA-Kme3). (**C**) Immunoprecipitated TAF10-HA was incubated with purified recombinant LOXL2 wild-type (rLOXL2 wt) or with catalytically inactive LOXL2 (rLOXL2 mut). After the LOXL2 reaction, samples were incubated with biotin hydrazide, and biotin-labelled proteins were pulled down with streptavidin beads. In parallel, after the LOXL2 reaction, samples were immunoprecipitated with anti-trimethylated lysine beads. oxTAF10 and TAF10-HA-Kme3 levels were determined by Western blot with an anti-HA antibody.

4. LOXL2 represses TFIID transcription-dependent genes by releasing TAF10 from target promoters

Since LOXL2 decreases TAF10 methylation, a modification that has been described to be required for the expression of a subset of TFIID-dependent genes[177], we checked whether genes regulated by this complex in a TAF10 methylation dependent manner, such as HOXA1, ERF1 and PLK1, are likewise regulated by LOXL2 oxidation of TAF10.

Using qRT-PCR, we observed that these genes were transcriptionally repressed when wild-type LOXL2 was infected in HEK293T cells. As a negative control, we checked the expression levels of HPRT and CCNE1, which are described to be independent of TAF10 methylation[177], and did not observe any changes in their transcriptional rates (*Figure R.6a*). Moreover, we demonstrated LOXL2 binding to those promoters, except CCNE1, by chromatin immunoprecipitation (ChIP) assays. LOXL2 binding was associated with a reduction in the levels of promoter-bound TAF10 (*Figure R.6b*), suggesting a role for LOXL2 in TAF10 release from these promoters when they are repressed. Importantly, when LOXL2 bound to the HPRT promoter, we did not observe TAF10 release; this is consistent with the fact that TAF10 is not methylated at this promoter and hence is not modified by LOXL2. In general, these results reinforce the idea of a

specific regulation of genes containing methylated TAF10 by LOXL2.

ChIP assays for RNAPII and TAF11 were then done to determine whether, in the affected promoters, TAF10 release also involved the lost of the other components of the TFIID complex and failed RNAPII recruitment. TAF11 is a subunit of the TFIID complex but is only incorporated in the holo-TFIID, the final form of the complex responsible for PIC formation (*Figure R.7*). The results showed less TAF11 and RNAPII enrichment, which could suggest that the whole TFIID complex is released and that RNAPII recruitment is impaired.

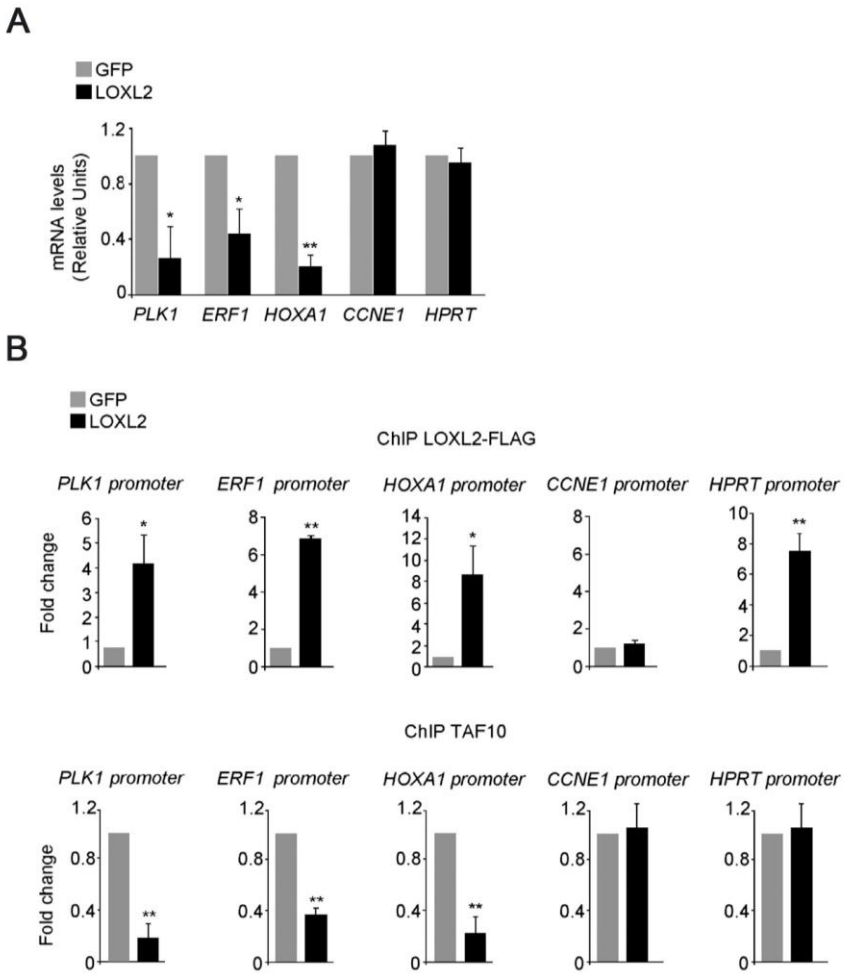

Figure R.6. LOXL2 represses TFIID-dependent genes by releasing TAF10 from target promoters. (**A**) Real-time quantitative RT-PCR (qRT-PCR) shows the changes in expression of PLK1, ERF1, HOXA1, CCNE1 and HPRT mRNA levels in HEK293T cells infected with either GFP or LOXL2 wt-FLAG. Gene expression was normalized against an endogenous control (Pumilio) and presented as RNA levels over those obtained in GFP-infected cells, which was set to 1. (**B**) LOXL2 wt-FLAG and endogenous TAF10 binding to those promoters was determined by ChIP assay in GFP-infected cells as compared to LOXL2 wt-FLAG infected cells. Data from real-time PCR (qPCR) amplifications of *PLK1*, *ERF1*, *HOXA1, CCNE1* and *HPRT* promoters were normalized to the input and expressed as fold-change relative to the data obtained in GFP-infected cells, which was set to 1. Error bars indicate standard deviation in at least three experiments. *$p < 0.05$; **$p < 0.01$.

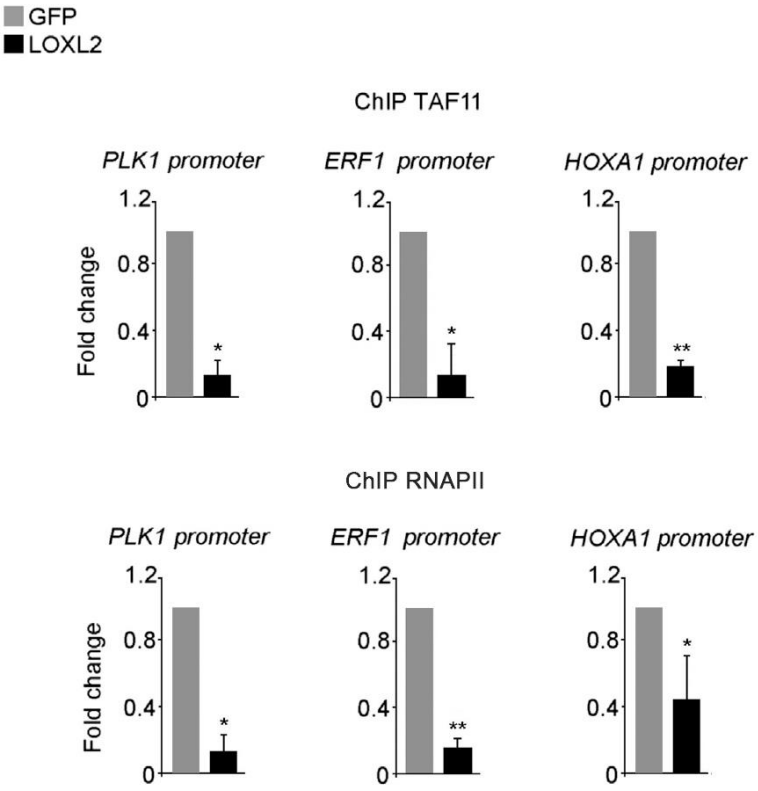

Figure R.7. TAF10 release affects TFIID and RNAPII binding to target promoters. Endogenous TAF11 and RNAPII binding to indicated promoters was determined by ChIP assay in GFP-infected cells as compared to LOXL2 wt-FLAG infected cells. Data from real-time PCR (qPCR) amplifications of *PLK1*, *ERF1* and *HOXA1* promoters were normalized to the input and expressed as fold change relative to the data obtained in GFP-infected cells, which was set to 1. Error bars indicate standard deviation in at least three experiments. *$p < 0.05$; **$p < 0.01$.

5. TAF10 oxidation by LOXL2 does not affect TFIID complex integrity

TFIID complex stability requires TAF10, as has been described in several studies[79,181]. To determine if TAF10 oxidation by LOXL2 affects TFIID complex integrity, we analysed the general composition of the complex in HEK293T cells expressing control vector, wild-type LOXL2 or mutant LOXL2. TFIID complex was immunoprecipitated from nuclear extracts using an anti-TBP antibody, and the presence of coimmunoprecipitated TAFs was determined by Western blot for the three conditions (*Figure R.8a*). We detected all the TAF proteins tested, indicating that LOXL2 overexpression does not alter the canonical composition of the TFIID complex in a general manner. However, when only oxidized TAF10 was immunoprecipitated, we observed that this modified TAF10 does not interact with any of the TBP and TAF subunits checked (*Figure R.8b*). This result indicates that although the general TFIID composition does not change, the subset of TAF10 that is oxidized may no longer be able to interact with the other members of the TFIID complex that we tested.

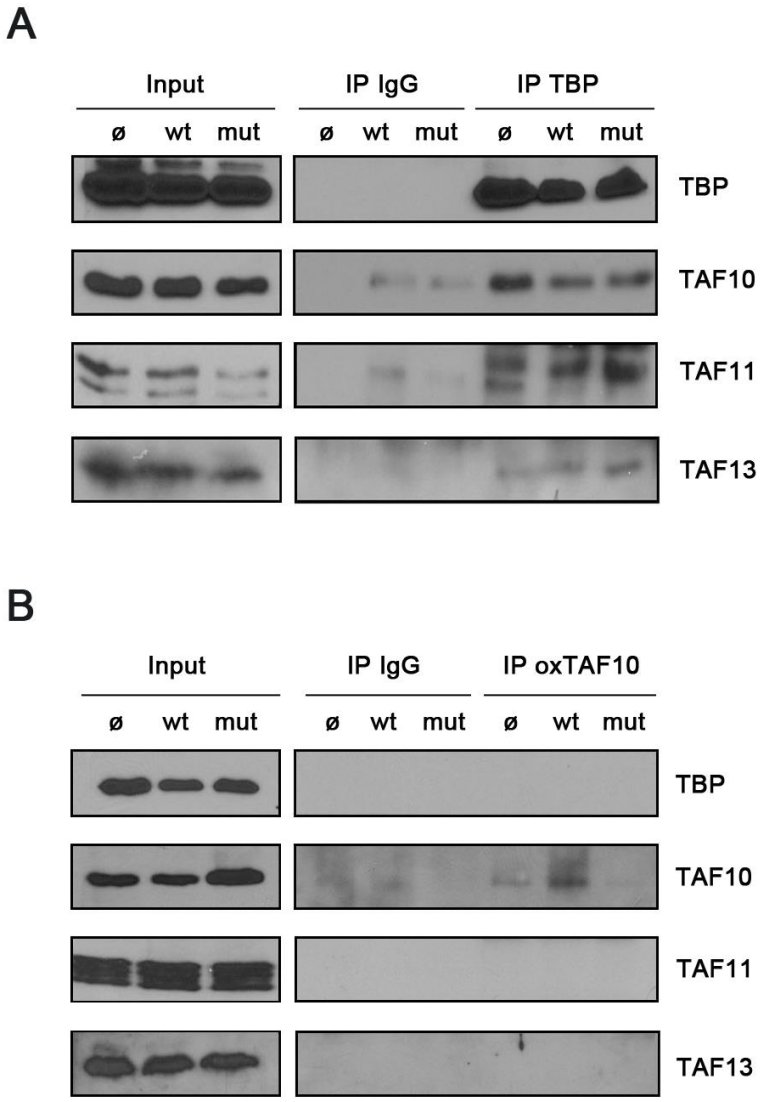

Figure R.8. General TFIID integrity is not affected upon TAF10 oxidation. (**A**) Nuclear extracts of HEK293T cells infected with GFP or LOXL2 wt-FLAG or LOXL2 mut-FLAG were immunoprecipitated with anti-TBP antibodies or irrelevant IgGs and analysed by Western blot with the indicated antibodies. (**B**) Nuclear extracts of HEK293T cells infected with GFP or LOXL2 wt-FLAG or LOXL2 mut-FLAG were incubated with biotin hydrazide, TAF10 was immunoprecipitated with an anti TAF10-antibody and then oxidized TAF10 was pulled down with streptavidin beads. The interactors were analysed by Western blot with the indicated antibodies.

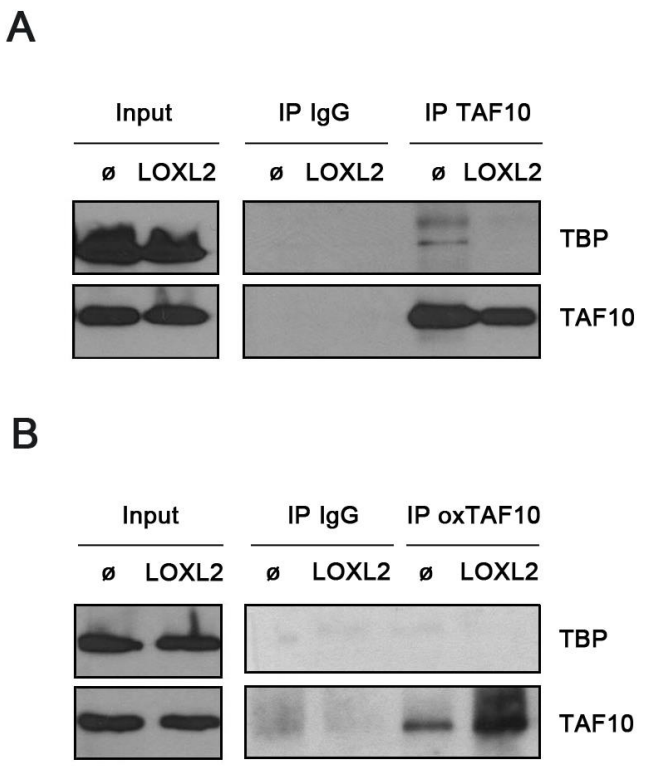

Figure R.9. Oxidized TAF10 does not interact with TBP. (**A**) Nuclear extracts of HEK293T cells infected with GFP or LOXL2 wt-FLAG were immunoprecipitated with anti-TAF10 antibodies or irrelevant IgGs and analysed by western blot with the indicated antibodies. (**B**) oxTAF10 was pulled down with streptavidin beads after a previous immunoprecipitation with an anti-TAF10 antibody from nuclear extracts incubated with biotin hydrazide. TBP interaction was analysed by Western blot.

Even though oxidized TAF10 did not show an interaction with other members of the complex, it was unclear if this was due to an actual lack of interaction or rather to a low immunoprecipitation efficiency. To answer this question, we immunoprecipitated comparable amounts of TAF10 and oxTAF10 and analysed for co-immunoprecipitating TBP

(*Figure R.9*). TBP was only detected in TAF10 immunoprecipitation (*Figure R.9a*), suggesting that TAF10 oxidation indeed impairs binding to other members of the complex. Moreover, when total TAF10 was immunoprecipitated in LOXL2 overexpressing conditions, the percentage of oxidized TAF10 increased and, as a consequence, the interaction between TAF10 and TBP was no longer detected (*Figure R.9a*).

6. LOXL2 oxidizes TAF10 and represses *Oct4*, *Nanog*, *Klf4* and *Sox2* transcription in ESCs, affecting the pluripotency state

TFIID is known to be required for the expression of pluripotency genes *Oct4*, *Nanog*, *Klf4*, and *Sox2* in mouse ESCs[80]. We therefore decided to check whether LOXL2 has a role as a negative regulator of the expression of these genes through TAF10 oxidation in ESCs.

We initially confirmed that mouse TAF10 is also trimethylated in this cellular mode. Mouse TAF10 was immunoprecipitated using two different antibodies and then re-immunoprecipitated with anti-trimethyl lysine antibodies (*Figure R.10a*). As expected, TAF10 was trimethylated, as in HEK293T cells.

Since TAF10 is modified in this model, we checked if it was also oxidized by LOXL2. Oxidized TAF10 was pulled down using streptavidin beads from nuclear extracts of ESCs infected with a control vector, wild-type LOXL2 or a catalytically inactive LOXL2 mutant (*Figure R.10b*). Indeed, cells overexpressing wild-type LOXL2 showed the highest levels of oxTAF10, in agreement with our results in HEK293T cells.

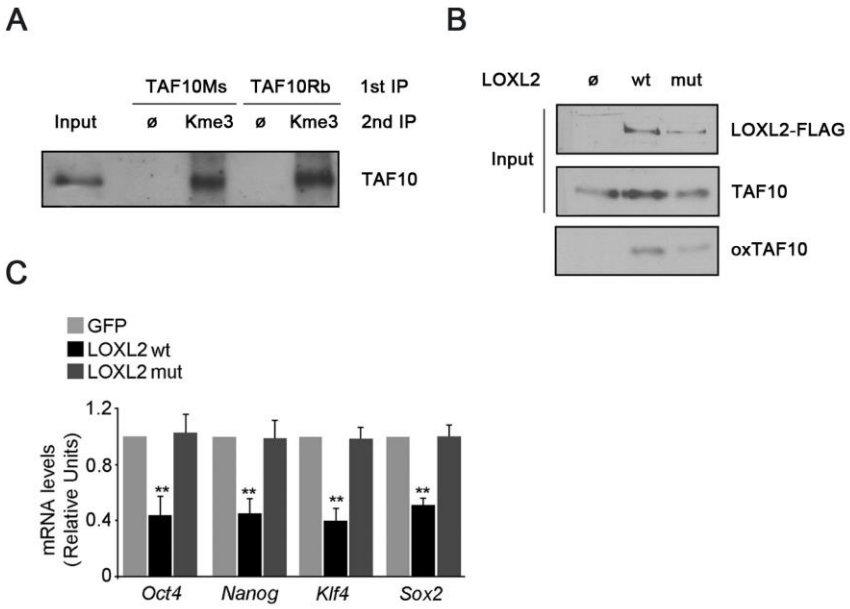

Figure R.10. LOXL2 oxidizes TAF10 in ESCs and represses pluripotency gene transcription. (**A**) TAF10 from ESCs was analysed by immunoprecipitation with two different antibodies and then re-immunoprecipitated with an anti-trimethyl lysine antibody. The levels of trimethylated TAF10 were analysed by Western blot against TAF10. (**B**) Cell lysates of ESCs infected with GFP, LOXL2 wt-FLAG, or LOXL2 mut-FLAG were incubated with biotin hydrazide, and oxidized proteins were precipitated with streptavidin beads. The oxidation levels of TAF10 were checked by western blot with an anti-TAF10 antibody. (**C**) qRT-PCR shows changes in expression of *Oct4*, *Nanog*, *Klf4* and *Sox2* mRNA levels in ESCs infected with GFP, LOXL2 wt-FLAG or LOXL2 mut-FLAG. Gene expression was normalized against an endogenous control (RPO) and presented as RNA levels over those obtained in GFP-infected cells, which was set to 1. RPO is the housekeeping gene ribosomal protein large, RPLP0.

Finally, we analysed expression levels of *Oct4*, *Nanog*, *Klf4* and *Sox2* pluripotency genes in ESCs infected with wild-type LOXL2 or with the catalytically inactive LOXL2. Interestingly, these genes showed lowered expression levels when the wild-type LOXL2 was expressed but not when the catalytically-inactive mutant was expressed (*Figure R.10c*).

These results indicated that active LOXL2 is a transcriptional repressor of these genes that are key to maintaining a pluripotent state in ESCs.

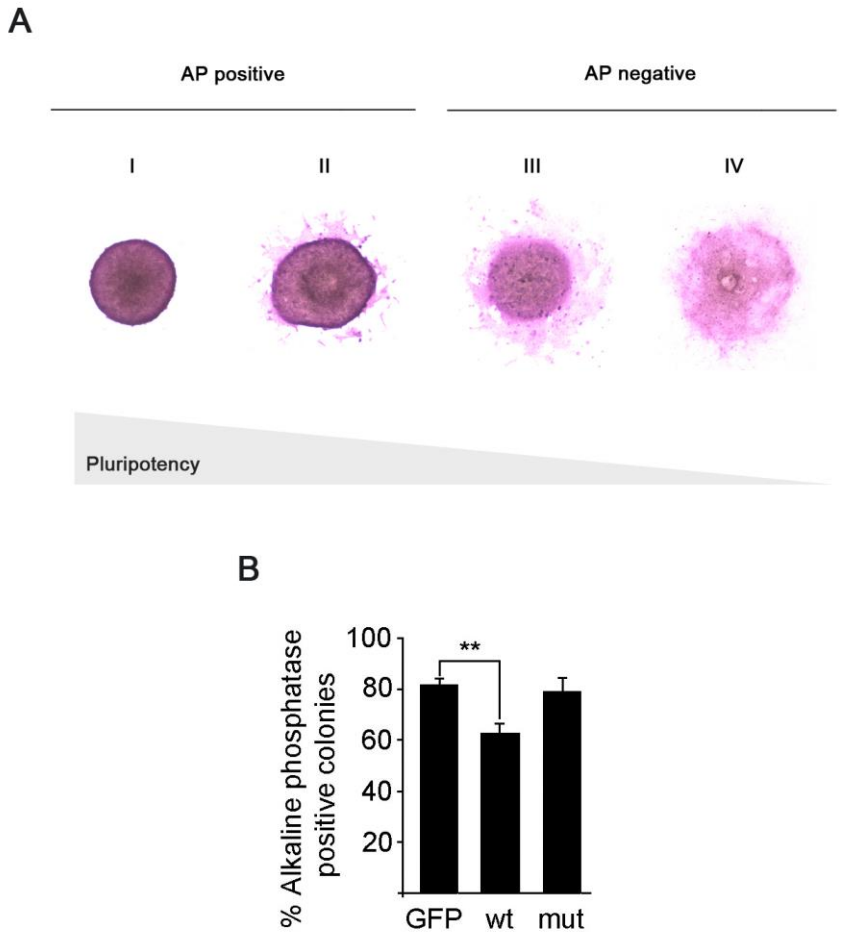

Figure R.11. LOXL2 overexpression decreases pluripotent state in ESCs. (**A**) Colony classification based on alkaline phosphatase (AP) staining and colony morphology for the quantification of AP positive or negative colonies. (**B**) Percentage of AP staining-positive ESC colonies that had been infected with GFP or LOXL2 wt-FLAG or LOXL2 mut-FLAG.

In agreement with a role for TFIID in maintaining pluripotent state[80], we also found that ectopic expression of wild-type LOXL2 but not the mutant form resulted in a differentiated morphology of ESCs, as evidenced by a decreased number of alkaline phosphatase–positive colonies that we used as a stem cell marker (*Figure R.11*).

Moreover, ChIP assays showed not only LOXL2 occupancy in those promoters, but also that this binding was concomitant with a release of TAF10 only when the wild-type LOXL2 was located in the promoters (*Figure R.12a*). Remarkably, we observed by re-ChIP assays that the remaining TAF10 is highly oxidized at the selected promoters in a LOXL2 catalytically-dependent manner (*Figure R.12b*).

In sum, these results suggested a tight regulation of the pluripotency state by LOXL2 in ESCs through TAF10 oxidation. Only active LOXL2 was able to oxidize TAF10 in pluripotency gene promoters, induce TAF10 release and, as a consequence, repress transcription of target genes under the control of those promoters.

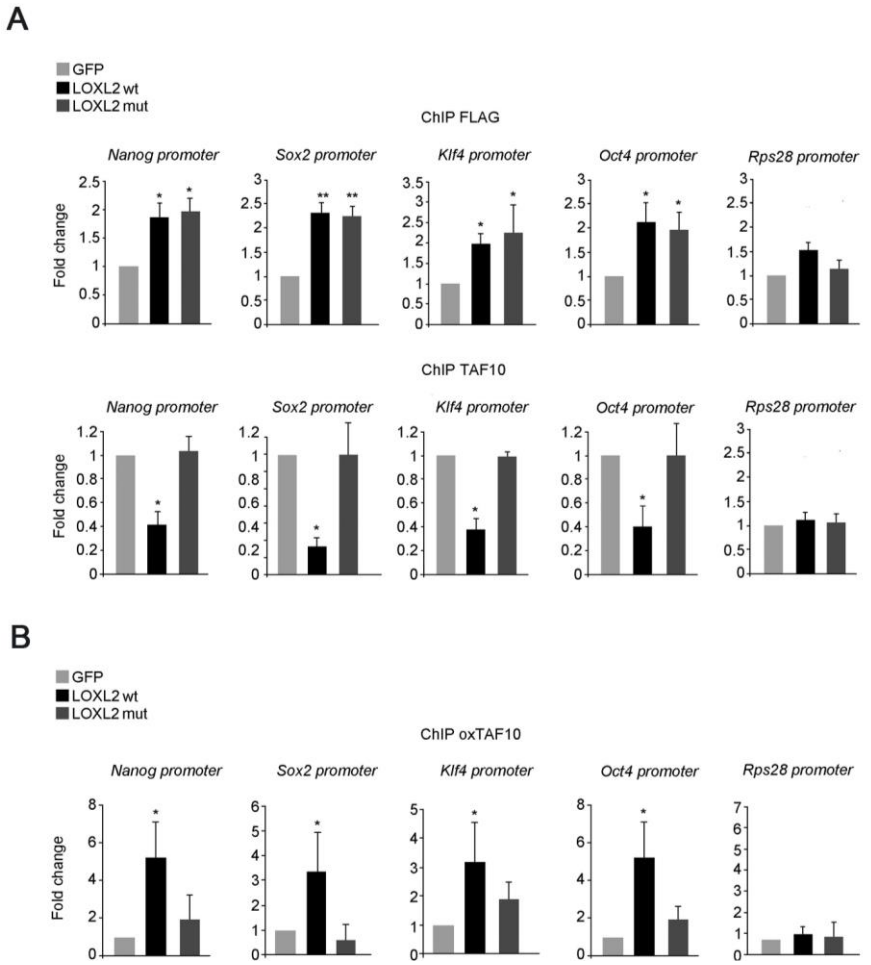

Figure R.12. LOXL2 binds to pluripotency gene promoters and induces TAF10 release in a catalytically-dependent manner. (**A**) LOXL2 wt-FLAG or mut-FLAG and endogenous TAF10 binding to target promoters was determined by ChIP in GFP-infected cells as compared to cells infected with either LOXL2 wt-FLAG or LOXL2 mut-FLAG. Data from qPCR of selected promoters was normalized to the input and expressed as fold-change relative to the data obtained in GFP-infected cells, which was set to 1. (**B**) TAF10 oxidation in the promoters was determined by re-ChIP in ESCs infected with GFP, LOXL2 wt-FLAG or mut-FLAG. Lysates incubated with biotin hydrazide were sequentially immunoprecipitated with anti-TAF10 and streptavidin beads. Data from qPCR was normalized to the total amount of immunoprecipitated TAF10 and to the input, and expressed as fold enrichment over the data obtained with GFP-infected cells. The value given for the GFP sample was set to 1. Error bars indicate standard deviation in at least three experiments. *$p < 0.05$; **$p < 0.01$.

7. Retinoic acid treatment induces LOXL2-dependent TAF10 oxidation, degradation and pluripotency gene repression

Since LOXL2 has been shown to be a repressor of key regulators of pluripotency network and to induce a decrease in the pluripotent state of ESCs, we further investigated its possible role during differentiation. For that purpose, we used an ESC differentiation model based on induction of neural progenitor cells by retinoic acid (RA) signalling.

We first checked LOXL2 expression levels in ES differentiating cells upon RA treatment at different time points and found that it was highly upregulated during this process. In concordance, we observed a decrease in the expression levels of pluripotency genes *Sox2*, *Nanog* and *Oct4*, as determined by qRT-PCR (*Figure R.13a*). Analysis of LOXL2 levels by Western blot showed that LOXL2 is expressed in basal conditions in ESCs and that, in accordance with the qRT-PCR, its levels increased after RA treatment (*Figure R.13b*). The oxidation levels of TAF10 also increased during this process, preceding a decrease in the global levels of this protein. The decrease in the TAF10 levels during RA-induced differentiation is probably due to changes in protein stabilization, since the transcriptional rates of TAF10 were maintained (*Figure R.13c*).

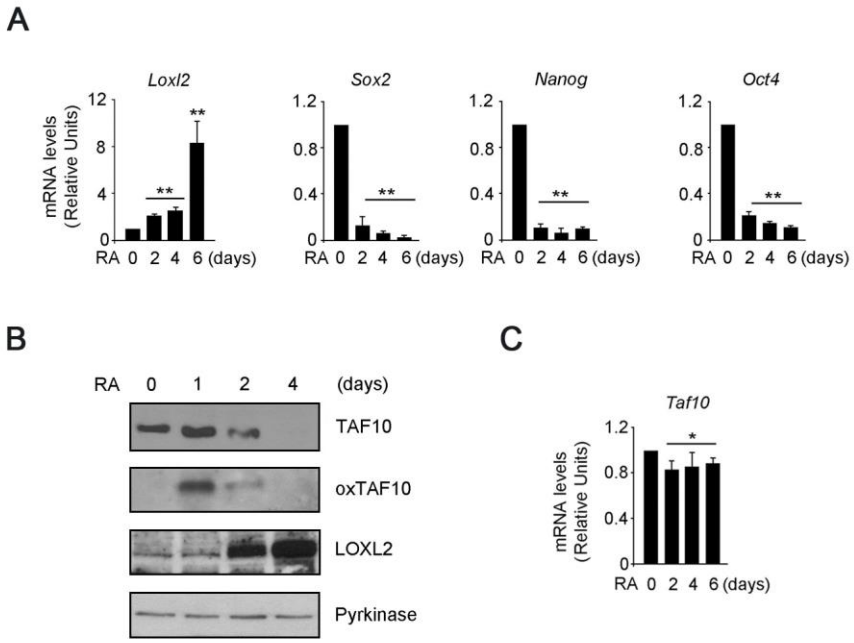

Figure R.13. LOXL2 is upregulated while TAF10 levels decrease upon RA treatment. (**A**) qRT-PCR shows the changes in expression of *Loxl2*, *Oct4*, *Nanog* and *Sox2* mRNA levels in ESCs treated with RA. (**B**) ESCs were treated with RA, and cell lysates were obtained at different time points, incubated with biotin hydrazide, and pulled down with streptavidin beads. The total and oxidized levels of TAF10 were checked by Western blot with an anti-TAF10 antibody. LOXL2 levels were analysed in the same cell lysates by Western blot with an anti-LOXL2 antibody. (**C**) qRT-PCR showed the changes in expression of *Taf10* mRNA levels in ESCs treated with RA. Gene expression was normalized against an endogenous control (RPO) and presented as RNA levels over those obtained without RA, which was set as 1. *$p < 0.05$; **$p < 0.01$.

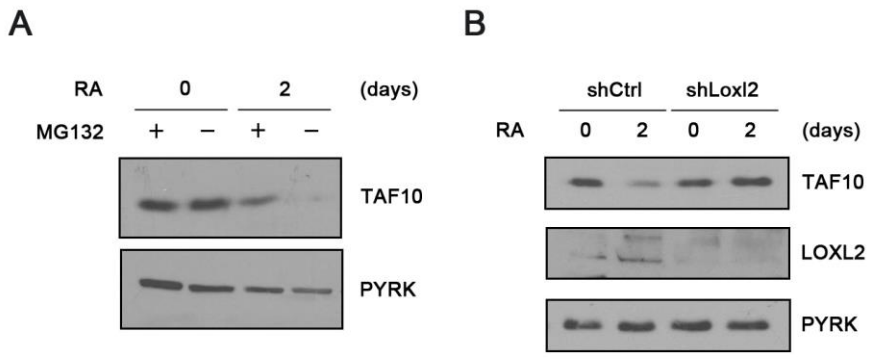

Figure R.14. Decreases in TAF10 levels are due to protein stability regulation and depend on LOXL2. (**A**) Western blot for TAF10 after 2 days of RA incubation and treatment with the proteasome inhibitor MG132 for 5 hours. (**B**) Western blot for TAF10 and LOXL2 after 2 days of RA treatment revealed the LOXL2-dependent degradation of TAF10 in ESCs infected with short-hairpin RNA control (shControl) or specific for LOXL2 (shLoxl2).

Decreases in total levels of TAF10 were blocked by the proteasome inhibitor MG132 (*Figure R.14a*). Furthermore, TAF10 degradation is LOXL2 dependent, since it was impaired when cells where infected with a specific shRNA against LOXL2 (*Figure R.14b*). To avoid side effects due to TAF10 degradation (which starts two days after RA treatment), all further experiments were done in ESCs after one day of RA.

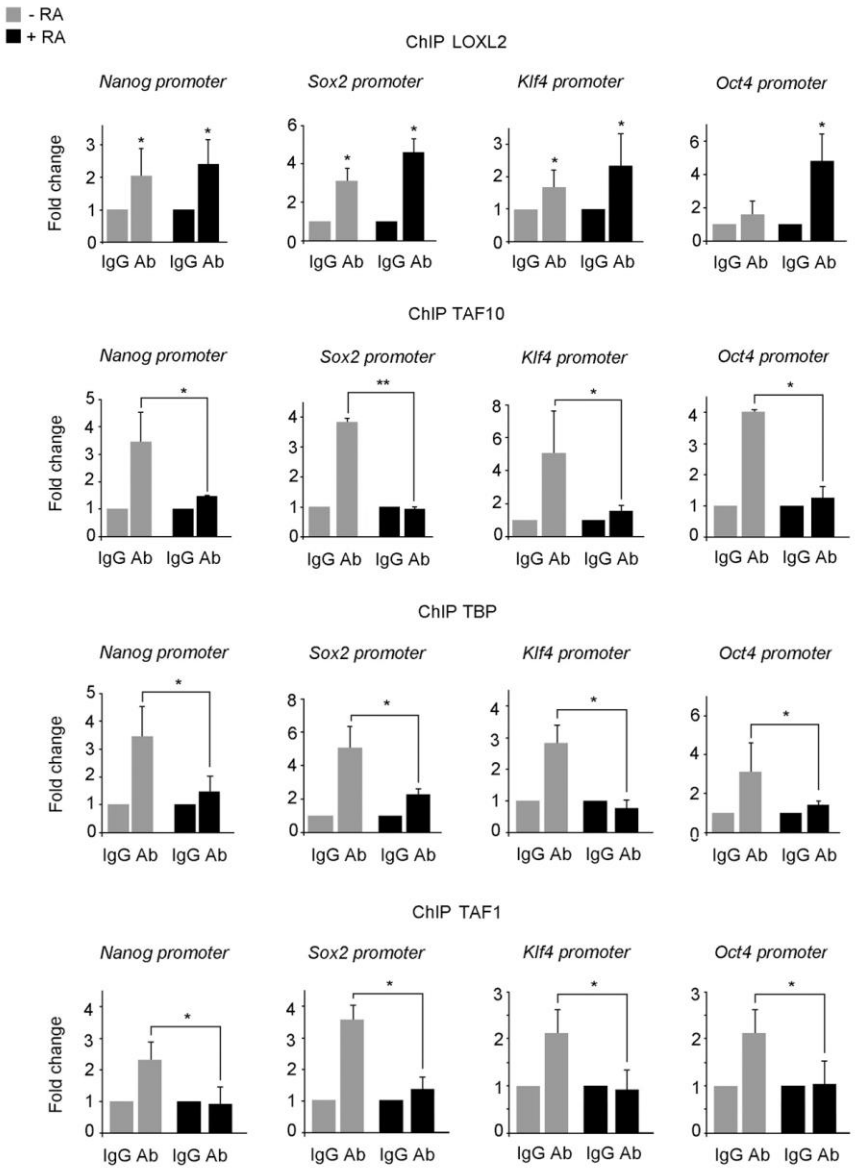

Figure R.15. RA treatment induces TAF10 and TFIID complex release from target promoters. Endogenous LOXL2, TAF10, TBP and TAF1 binding to *Oct4*, *Nanog*, *Klf4*, and *Sox2* promoters was analysed by ChIP after one day of RA treatment. Data from qPCR were normalized to the input and expressed as fold change relative to the data obtained with irrelevant IgG, which was set to 1. Error bars indicate standard deviation in at least three experiments.

Since we had shown ectopic LOXL2 binding and TAF10 release in the pluriporency gene promoters in ESCs, we next analysed whether endogenous LOXL2 was bound to these promoters, inducing TAF10 release in RA differentiating ESCs. By ChIP assays, we observed endogenous LOXL2 and TAF10 bound to *Nanog*, *Sox2*, *Klf4*, and *Oct4* promoters. After one day of treatment, TAF10 was released from the tested promoters, together with TBP and TAF1, two members of the TFIID complex (*Figure R.15*).

This release was promoter specific, since none of the additional sequences that we checked for LOXL2, TAF10, TBP, or TAF1 in a control promoter or intra- and intergenic regions showed the same behaviour (*Figure R.16*).

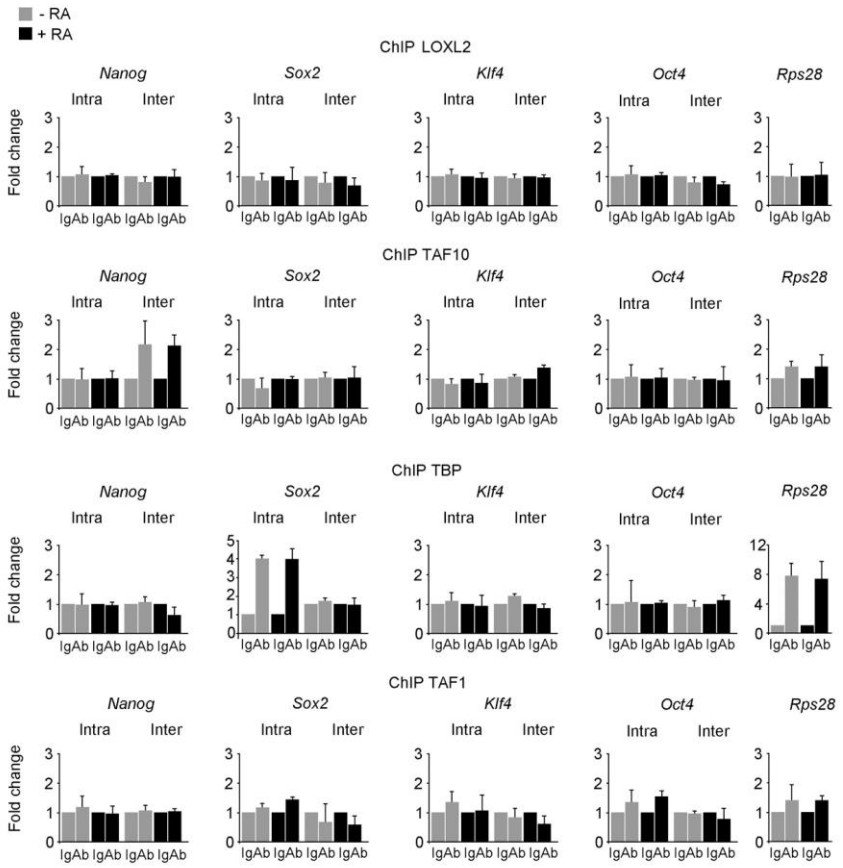

Figure R.16. RA effects observed in pluripotency gene promoters are site specific. Endogenous LOXL2, TAF10, TBP and TAF1 binding was analysed by ChIP to *Oct4*, *Nanog*, *Klf4*, and *Sox2* intergenic (inter) and intragenic (intra) control regions and *Rps28* promoter in ESCs treated for one day with RA. Data from qPCR were normalized to the input and expressed as fold-change relative to the data obtained with irrelevant IgG, which was set as 1. Error bars indicate standard deviation in at least three experiments.

Again, we observed by ChIP that the remaining TAF10 is highly oxidized in the selected promoters after one day of RA treatment (*Figure R.17a*).

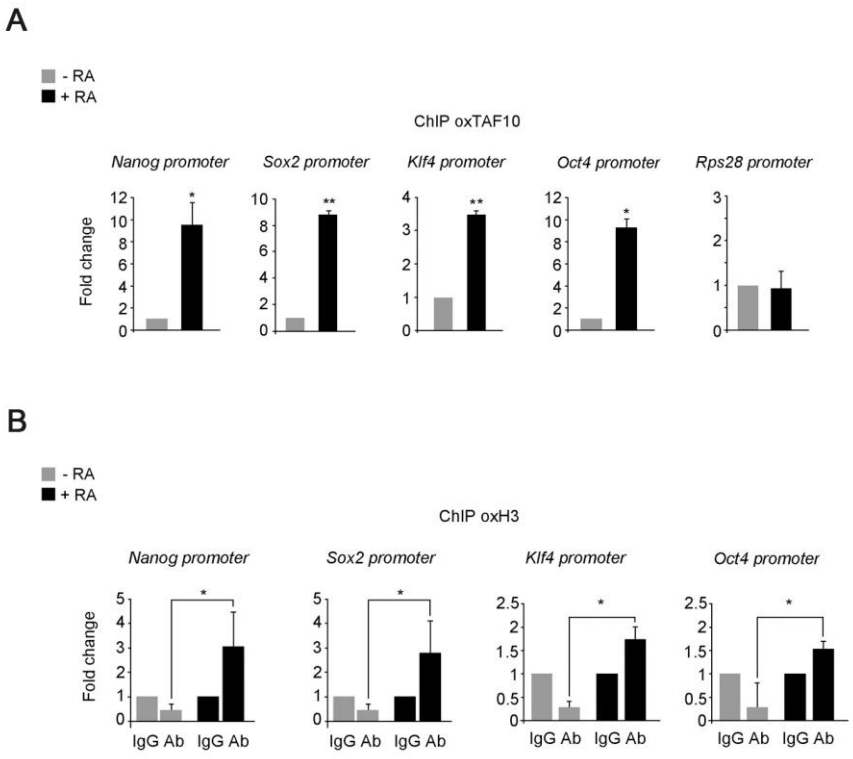

Figure R.17. RA treatment leaves oxidized forms of TAF10 and histone H3 in pluripotency gene promoters. (**A**) TAF10 oxidation in the selected promoters was determined by re-ChIP in ESCs that were untreated or after one day of RA treatment. The lysate was incubated with biotin hydrazide before ChIP. Extracts were sequentially immunoprecipitated with anti-TAF10 and streptavidin beads. DNA binding was quantified by qPCR. Data were normalized to the total amount of immunoprecipitated TAF10 and to the input, and expressed as fold enrichment over the data obtained in nontreated cells. The value given for the nontreated cells samples was set to 1. (**B**) oxH3 was determined by re-ChIP in the indicated promoters using biotin hydrazide and sequential immunoprecipitation of histone H3 and pull down with streptavidin beads. Data were normalized to the total amount of immunoprecipitated histone H3 and to the input, and expressed as fold enrichment over the data obtained with irrelevant IgG. The value given was set as 1. *$p < 0.05$; **$p < 0.01$.

LOXL2 binding to these promoters was also associated with increased histone H3 oxidation levels, which is a known specific LOXL2 substrate[91]. Although LOXL2 is already present in these target promoters at basal conditions, histone H3 oxidation was only observed after RA treatment, suggesting that LOXL2 was fully activated by RA (*Figure R.17b*).

Thus, we observed LOXL2 bound to pluripotency gene promoters, resulting in TAF10 oxidation and release, together with the TFIID complex, upon RA treatment. Moreover, histone H3 was oxidized in the same promoters, resulting in an epigenetic mark for transcription repression in these regions. Since all these events should lead to transcription repression of target genes, we further investigated if pluripotency gene repression occurring upon RA differentiation was LOXL2 dependent.

We generated LOXL2-deficient ESCs using a specific short-hairpin RNA and then induced differentiation through RA treatment. When comparing shLoxl2 to shControl cells, *Nanog*, *Klf4*, *Sox2* and *Oct4* maintained higher expression levels after 1 day of RA treatment (*Figure R.18a*). This result suggested that LOXL2 expression was required for a proper downregulation of pluripotency genes in this differentiation pathway, as when it was depleted, genes were not repressed to the same level as in control cells.

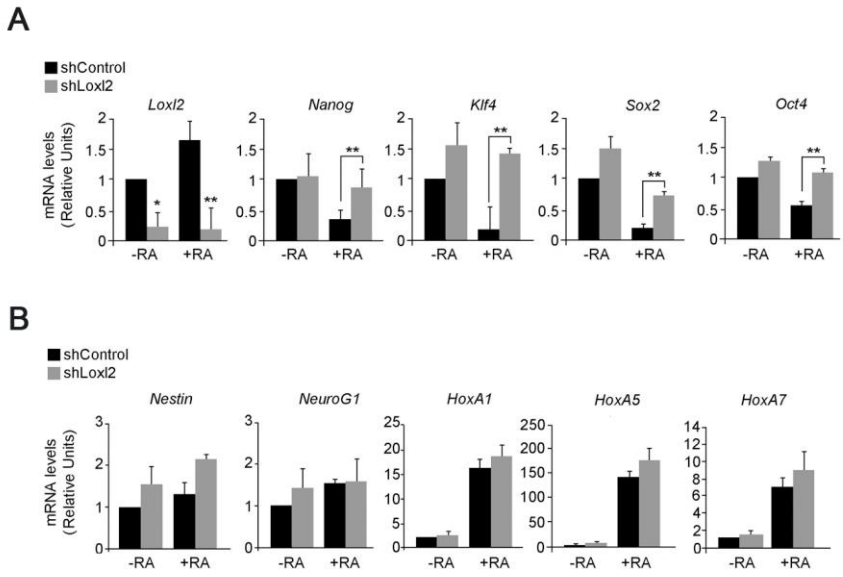

Figure R.18. Pluripotency gene repression after RA treatment is impaired in the absence of LOXL2. (**A**) Changes in expression of *Loxl2*, *Oct4*, *Nanog*, *Klf4* and *Sox2* mRNA levels in ESCs infected with shControl or shLoxl2 and treated with RA were determined by qRT-PCR. Gene expression was normalized against an endogenous control (RPO) and presented as RNA levels over those obtained in shControl-infected cells at day 0, which was set to 1. (**B**) qRT-PCR shows changes in expression of *Nestin*, *NeuroG1*, *HoxA1*, *HoxA5* and *HoxA7* mRNA levels in ESCs treated with RA. Gene expression was normalized against an endogenous control (RPO) and presented as RNA levels over those obtained in shControl-infected cells at day 0, which was set to 1. *p < 0.05; **p < 0.01.

To check if LOXL2 depletion also affects the induction of differentiation related expression pattern, we analysed the expression of *Nestin* and *NeuroG1* genes, which are markers for neural progenitors and neuronal differentiation[156,182], and three different *HoxA* genes (*HoxA1*, *HoxA3* and *HoxA5*), which are development related genes rapidly upregulated upon differentiation[183,184]. Importantly, the lack of LOXL2 did not affect the induction of differentiation-related genes in

ESCs upon RA treatment, suggesting a specific function for LOXL2 in pluripotency network repression rather than for activation of a differentiation expression pattern (*Figure R.18b*).

Nevertheless, LOXL2-deficient cells started to die with RA treatment, evidenced by decreases in cell number and the presence of cells with senescent-like morphology in cells treated with shLoxl2 as compared to those with the shControl (*Figure R.19a*). This result was further confirmed by clonogenic assay of ESCs infected with shControl or shLoxl2 and treated with RA: shLoxl2 cells showed a significant reduction in cell viability following RA treatment (*Figure R.19b*).

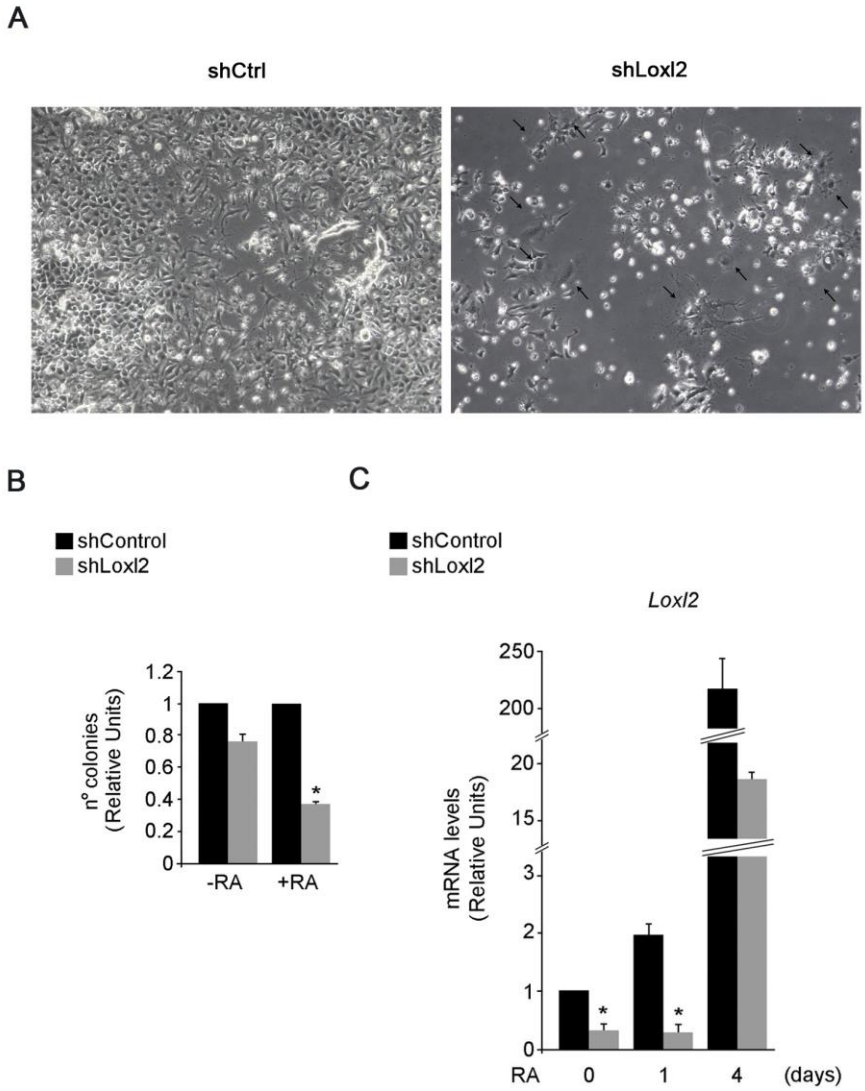

Figure R.19. RA treatment reduces viability of shLoxl2 ESCs. (**A**) LOXL2 knockdown reduced cell viability and induced senescent-like phenotypical changes (marked with arrows) upon RA treatment in ESCs. (**B**) Cell survival was analysed by a clonogenic assay by seeding shControl and shLoxl2 ESCs at clonogenic densities and treating them with RA. shLoxl2-infected cells colony numbers were normalized over shControl-infected cells in each. (**C**) Changes in *Loxl2* expression in ESCs infected with shControl or shLoxl2 and treated with RA were determined by qRT-PCR. Gene expression was normalized against an endogenous control (RPO) and presented as mRNA levels over those in shControl-infected cells at day 0, which was set to 1. *$p < 0.05$.

Interestingly, despite any reduction in cell viability, some shLoxl2 cells survived RA treatment and did differentiate. However, when analysing *Loxl2* expression levels, we observed that shLoxl2 cells that had survived RA incubation and differentiated had also a recovered expression of *Loxl2* (*Figure R.19c*), indicating that LOXL2 expression is indeed required in ESCs for a proper exit from pluripotency and correct differentiation.

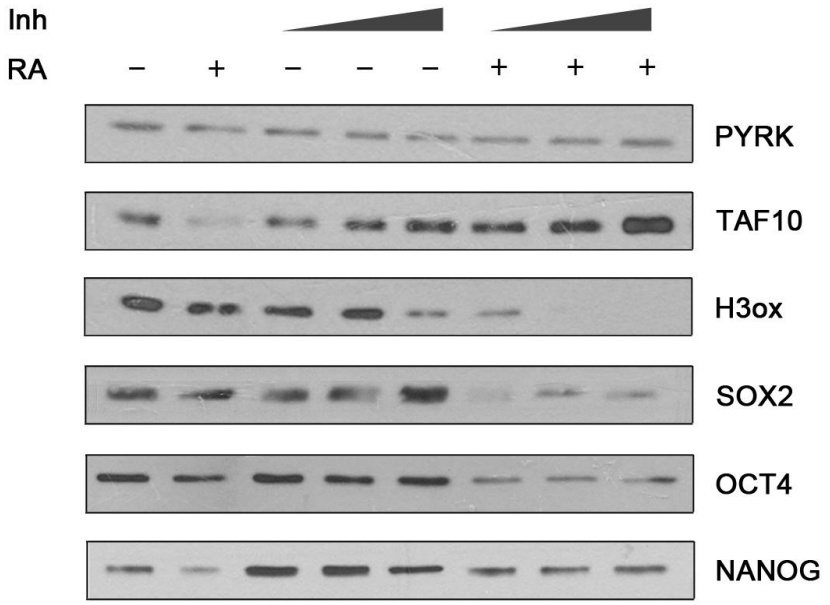

Figure R.20. Inhibition of LOXL2 activity increases pluripotency related protein levels. ESCs were treated with increasing concentrations of LOXL2 inhibitor (10 μM, 20 μM and 50 μM) without or with RA. Total cell extracts were obtained, and the indicated proteins levels were analysed by western blot.

To determine if LOXL2 dependent repression of pluripotency genes depended on its catalytic activity, we used a LOXL2 inhibitor to treat cells upon RA treatment and checked expression of a subset of pluripotency-related proteins (*Figure R.20*). As a control, we verified that levels of oxidized histone H3 decreased while TAF10 protein levels increased upon inhibitor treatment. In both cases, the effect of the inhibitor was stronger upon RA treatment, indicating again that RA could be fully activating LOXL2.

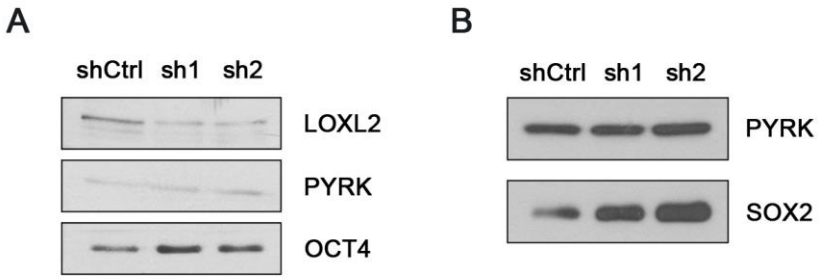

Figure R.21. LOXL2 depletion induces expression of pluripotency factors in ESCs. Total extracts of ESCs infected with a control short-hairpin RNA or two different specific short-hairpin RNAs for *Loxl2* were analysed by Western blot for the indicated antibodies.

We observed a tendency for higher protein levels for pluripotency-related proteins in both RA-treated and non-treated cells. Hence, results indicate that basal expression levels of LOXL2 already have effects on ESCs. In fact, knocking down LOXL2 expression with two specific short-hairpin RNAs resulted in similar effects (*Figure R.21*),

reinforcing the idea that LOXL2 has a role in the equilibrium of pluripotency dynamics in ESCs.

It is likely that the increase in LOXL2 activity after RA treatment is due to the dual effect of a fully activated LOXL2 on one hand and higher expression levels of LOXL2 on the other.

8. Global transcriptome analysis implicates LOXL2-dependent genes in neural progenitor differentiation

Considering the critical role for LOXL2 in RA-induced ESC differentiation, we further investigated the scope of LOXL2 regulation by performing a microarray analysis on LOXL2-depleted ESCs that had been treated or not with RA.

The comparison between the transcriptomes after RA treatment of shLoxl2 against the shControl revealed 524 genes that were regulated only in control conditions upon RA treatment but not in the absence of LOXL2 (*Figure R.22a*). These genes all depend on LOXL2 for repression upon RA treatment, either directly or indirectly. In concordance, GO analysis of these genes showed association with embryonic development and function. The observed changes were validated by RT-qPCR for selected genes (*Figure R.22b*).

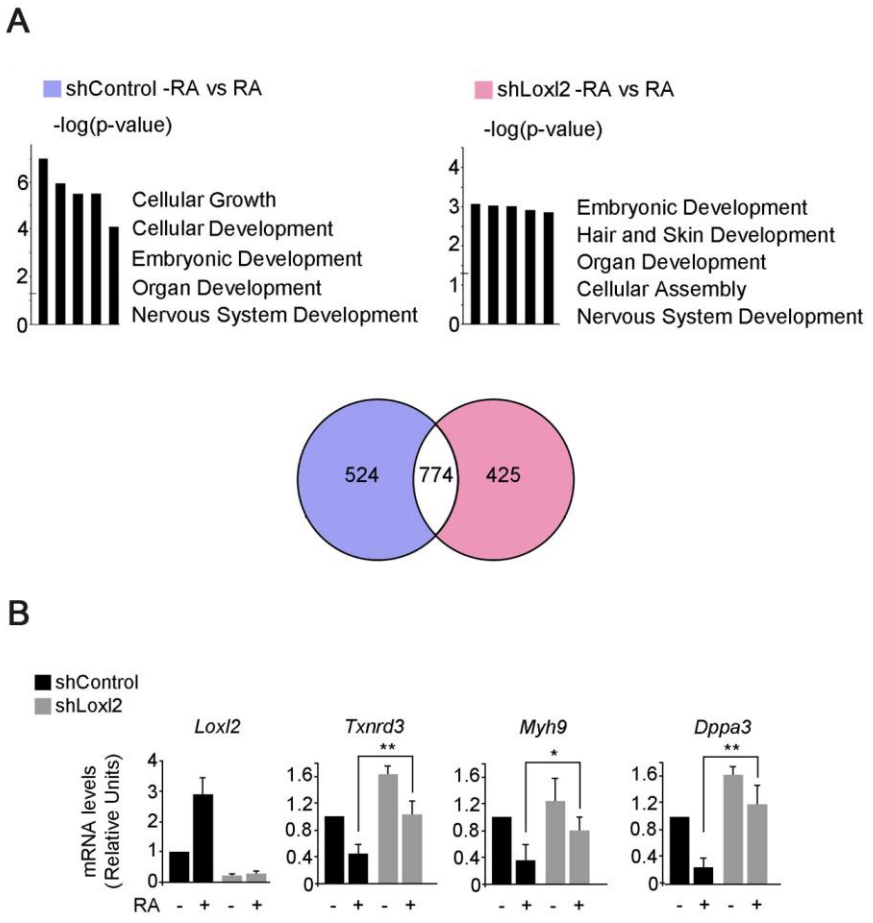

Figure R.22. Global transcriptome analysis of LOXL2-dependent genes in ESC differentiation. (**A**) Venn diagram representing data obtained from the microarray analysis of gene expression profiles in ESCs shControl and shLoxl2 treated 1 day with RA. Gene ontology (GO) analyses of genes only deregulated in shControl (purple) and shLoxl2 (pink) conditions after 1 day RA treatment, using the gene annotation tool from the Ingenuity database, are shown. (**B**) Validation by RT-qPCR of selected LOXL2-dependent genes. Results are from three independent experiments.

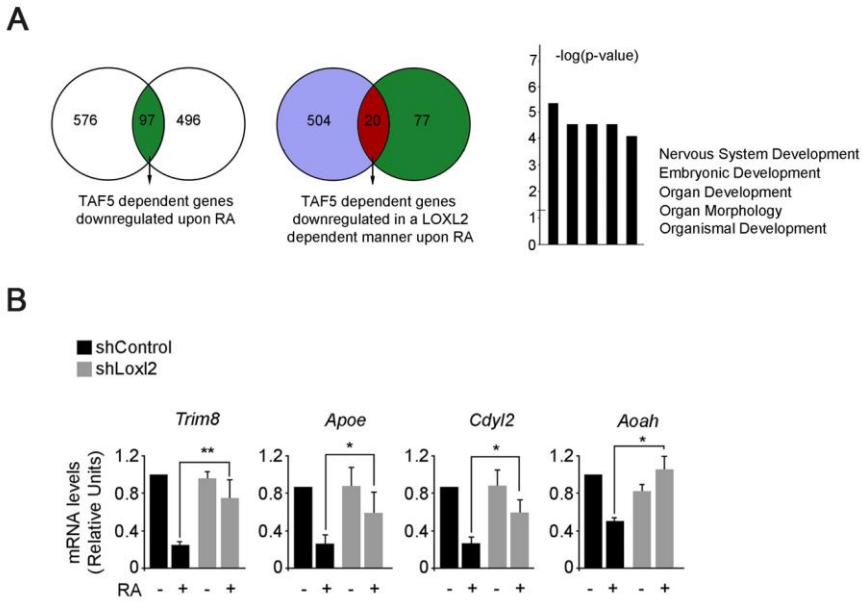

Figure R.23. Global transcriptome analysis confirms that LOXL2 has a role in embryonic and neural system development. (**A**) Left, Venn diagram representing gene expression profile data from the knockdown of TAF5 in ESCs (obtained form the GEO under accession number GSE33644), and the gene expression profile in ESCs shControl treated with RA for 1 day. Genes shown in green are those that are TAF5 dependent and downregulated upon RA treatment. Right, Venn diagram representing gene expression profile of TAF5-dependent genes downregulated after RA treatment (green), and gene expression profiles of genes deregulated in shControl but not in shLoxl2 conditions, together with GO analysis of this set of genes. (**B**) Validation by RT-qPCR of selected genes, in three independent experiments *p < 0.05; **p < 0.01.

To determine which genes that are LOXL2-dependent upon RA induction are also TFIID-dependent in ESCs, we compared the gene expression profile of the available data for TAF5 knockdown in ESCs[80]. We considered genes that were downregulated in the absence of TAF5 to be TFIID transcription–dependent genes and then used our transcriptome data in control ESCs treated with RA to

determine which TFIID-dependent genes were downregulated during differentiation; this resulted in 97 genes (*Figure R.23a*).

Finally, we crossed this 97-gene data set with the list of genes that did not change with RA in the absence of LOXL2 (524 genes; *Figure R.23a*). With this analysis, we obtained 20 TFIID-dependent genes that were downregulated during differentiation in a LOXL2-dependent manner (*Table A.5*), which represents 20% of TFIID-dependent genes in RA differentiation. Interestingly, GO analyses of these genes corresponded with embryonic development and nervous system development. Again, the observed changes for selected genes were validated by RT-qPCR (*Figure R.23b*).

9. LOXL2 regulates *Sox2* expression and neural fate specification in zebrafish

As our results indicate a role for LOXL2 in embryonic development in cell culture conditions, we addressed the possible functions of LOXL2 in balancing pluripotency and cell differentiation in an *in vivo* model using zebrafish embryos. The zebrafish genome encodes eight lysyl oxidases in total, orthologous of the four human lysyl oxidases, since the zebrafish genome underwent a second round of genome duplication not present in other nonteleost vertebrates[185]. Based on nucleotide sequence alignment, the human *LOXL2* gene is orthologous to the zebrafish *Loxl2a* and *Loxl2b* genes, which share a similar percentage of amino acid sequence with human LOXL2 protein (*Figure R.24*; upper panel).

Therefore, we based our functional analysis on these two lysyl oxidases. At 24 hours postfertilization (hpf), *Loxl2a* was expressed throughout the embryo, with high staining in the eyes, the mesencephalon and the hindbrain boundaries (*Figure R.24*; lower panel, arrowheads). A similar but not identical pattern of expression was observed for *Loxl2b*, with ubiquitous expression throughout the embryo.

	LOXL2_HUMAN	LOL2A_DANRE	LOL2B_DANRE
LOXL2_HUMAN	100	67.57	70.39
LOL2A_DANRE	67.57	100	72.85
LOL2B_DANRE	70.39	72.85	100

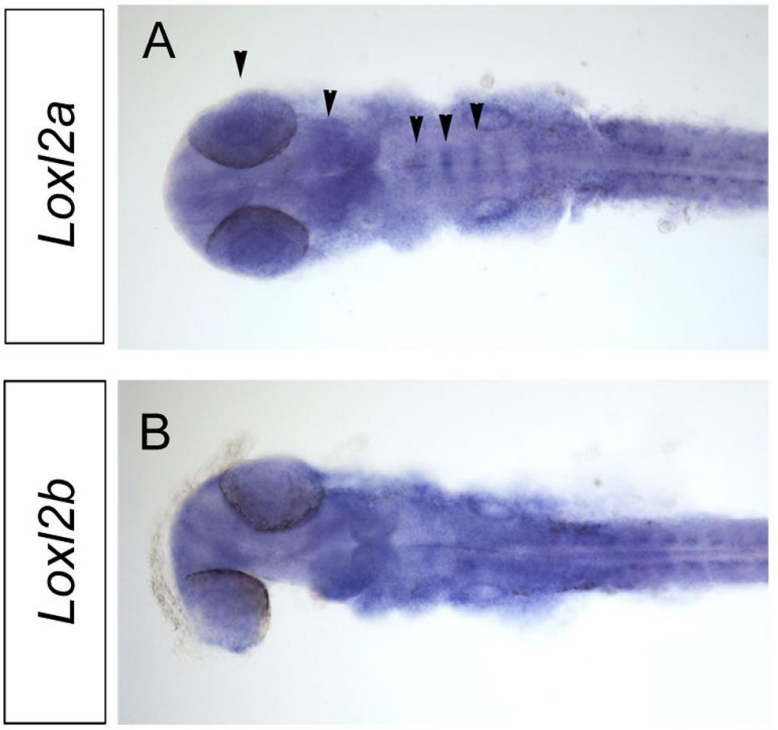

Figure R.24. *loxl2a* and *loxl2b* are expressed in 24 hpf zebrafish embryos. Upper panel, identity matrix of aminoacid sequences of human LOXL2 and two orthologous zebrafish LOL2A and LOL2B proteins. Lower panel, *loxl2a* is expressed throughout the embryo; the high-staining regions indicated with arrowheads were the eyes, mesencephalon and hindbrain boundaries. *loxl2b* presents an ubiquitous expression pattern throughout the embryo.

Sox2 is one of the core components of the pluripotency gene network that, in vertebrates, is involved initially in epiblast development and later in neural development[186]. During the later stages, SOX2 maintains neural stem cells in an undifferentiated state by repressing proneural genes *Neurogenin1* and *NeuroD*[187].

We tested whether, as in ESCs, LOXL2 is involved in the transcriptional control of *Sox2 in vivo* by inhibiting LOXL2 orthologs in zebrafish embryos (*Figure R.25*). In embryos injected with Loxl2a and Loxl2b morpholinos, *Sox2* expression was upregulated in the central nervous system (CNS), and in particular, in the eye, hindbrain and spinal cord (*Figure R.25A and R.25B*, see arrows).

Concomitant with the induction of *Sox2* expression, the levels of *NeuroD*-positive cells (e.g., cells entering neuronal differentiation) were reduced in the domains mentioned above (*Figure R.25C-R.25F*; see insets of spinal cord). These alterations were associated with morphological defects of the brain; in particular, the anterior brain was rounder and the eyes presented a flatter curvature (compare images *Figure R.25C'* with *Figure R.25D'*, and *Figure R.25E'* with *Figure R.25F'*).

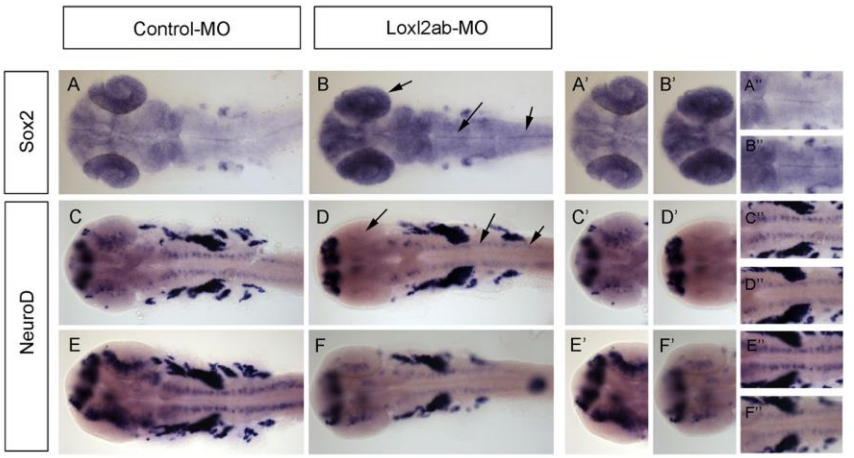

Figure R.25. LOXL2 regulates *Sox2* expression and neural fate specification in zebrafish. In loxl2ab-morpholino (MO)-injected embryos, *Sox2* expression increased (B) (n = 9/12) as compared with control (A) (n = 16/16), in particular in the eye, hindbrain, and spinal cord (B; arrows); see insets of eyes (A' and B') and spinal cord (A'' and B''). *NeuroD* expression was reduced in the same domains in loxl2ab-MO-injected embryos (D and F) (n = 12/16; arrows) as compared to the controls (C and E) (n = 10/10); see insets of eyes (C'-F') and spinal cord (C''-F'').

Finally, we also found that inhibition of *Loxl2a* and *Loxl2b*, validated by PCR for *loxl2a* (*Figure R.26A*) while *Loxl2b* morpholino was already described[185], compromised embryo development. Only 33.5% of injected embryos survived as compared to 44.5% of control-injected embryos (*Figure R.26B*).

Since *loxl2a* and *loxl2b* depletion is not complete, and since other members of the lysyl oxidase family could be compensating this depletion, the fact that survival was reduced suggests a key role for LOXL2 in embryonic development.

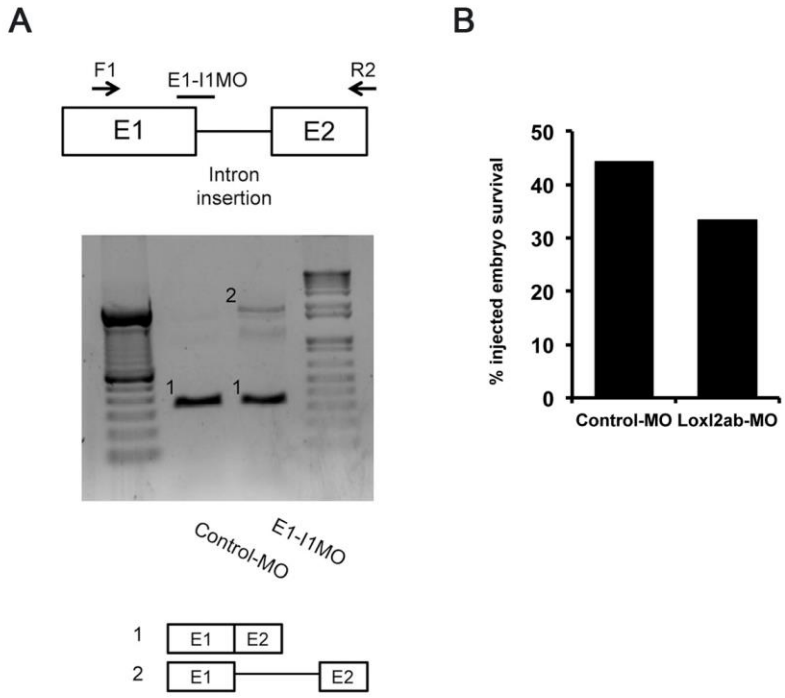

Figure R.26. Depletion of *loxl2a* and *loxl2b* by morpholinos reduces embryo survival. (**A**) Schematic representation of a region of zebrafish *loxl2a* pre-mRNA. Exons (E) are depicted in squares and introns (I) as lines. The bar indicates the location of the splice site targeted by the *loxl2a*-MO (E1-I1MO). Arrows indicate the location of primers used to amplify the cDNA in the PCR. The ATG is located in exon 1. The intron 1 (1285 b) is inserted by the E1-I1MO injection leading to a frame shift and a premature stop codon. RT-PCR from 18 hpf embryos infected with control-MO or E1-I1MO. Primers F1-R2 were used to check the effect on splicing of the *loxl2a* gene of the E1-I1MO. Numbers in the gel correspond to drawings below depicting the exon-intron composition of every PCR products. (**B**) Percentage of surviving injected embryos. Embryos were injected with either Control-MO or Loxl2ab-MO. Surviving rates for each condition are represented as percentage of the total embryos injected.

DISCUSSION

Histone-modifying enzymes are known to be able to also modify non-histone substrates[188-190]. Here we demonstrate that the histone-modifying enzyme LOXL2 is responsible for TAF10 transcription factor oxidation, which is a subunit of TFIID general transcription factor complex, causing the repression of a subset of TFIID-dependent genes. We demonstrate that in ESCs, this repression is key for maintaining the balance between pluripotency and differentiation.

1. LOXL2 new putative substrates

An unbiased proteomic approach was performed to detect putative non-histone substrates using nuclei enriched HEK293T extracts. We obtained 117 candidates, for which string analysis showed enrichment of four different modules: TFIID complex, ribosomal proteins, splicing factors and actin related proteins. However, when analysing the involved cellular pathways, several substrates were related not only to basal transcription factors but also to DNA damage repair and metabolism. It is worth noting that mass spectrometry analysis has several detection limitations, and that some proteins may cover the signal of others because they are more abundant or due to their biochemical properties. Further, some peptides ionize easier, fly better and fractionate in more interpretable patterns than others, making

their characterization easier. Thus, some proteins are easier to detect than others due to their sequence and biochemical properties, and these can cover other proteins, which are then more difficult to detect even if they are more abundant than the easily-detected protein.

Although they are enriched in interactomes, actin-related proteins and ribosomal proteins are very prone to be non-specific candidates, since they are abundant proteins commonly found as contaminants in affinity purification and mass spectrometry data[191]. However, a role for LOXL2 in these functions would not be surprising. First, LOXL2 is a regulator of the EMT process, in which cells redistribute fibres to reduce adherent junctions and acquire motility[102]. Moreover, a role for LOXL2 in elastin and collagen fibres oxidation that reorganizes their structure is already described[192,193]. Hence, LOXL2 activity towards actin fibre reorganization could be important during EMT.

Although ribosomal proteins are common contaminants, we cannot overlook them as putative candidates. When analysing putative partners of another LOX family member, LOXL1, we found the histone modifying enzyme FBRL, which specifically binds to ribosomic DNA repeats, to be highly enriched together with other nucleolar proteins where ribosomal proteins are found (unpublished data). Even if this is not the same pattern that was found when looking at

LOXL2 putative substrates, ribosomal proteins could potentially be LOXL2 substrates. We also obtained splicing related factors, which could be indirectly immunoprecipitated together with histones, since in this experiment histone H3 and histone H1 were significantly enriched. In any case, further research and differently designed experiments including controls to discard false positives would be required to distinguish if these are real candidates or not.

Among the more confident candidates, we found histone H1 as putative substrate. Histone H1 is the substrate for which we found more predicted lysines suitable for SET7/9 methylation, the enzyme responsible for H3K4 methylation. Moreover, several methylations have already been described for histone H1[194]. Although the function of most of these PTMs are still unknown, H1.4K26 dimethylation is for example described to recruit HP1α leading to transcriptional repression[195]. Our group recently described that LOXL2 activity induces release of HP1α from pericentromeric regions during EMT. This effect is due to a transcriptional downregulation of major satellite transcripts controlled by oxidation of H3. It has been described that pericentromeric transcripts help HP1α binding to pericentromeric heterochromatin[196], therefore a reduction in the pericentromeric transcripts would participate in HP1α release. Moreover, if histone H1 methylation is important for HP1α binding, oxidation by LOXL2 could impair this interaction, in

agreement with a LOXL2 function in impairing HP1α binding to chromatin.

Interestingly, histone H1 has already been proposed as a candidate substrate for the LOX enzyme[197], as when these are incubated together *in vitro*, hydrogen peroxidase is produced. Moreover, their interaction *in vitro* has also been demonstrated[198]. LOX overexpression is related to the regulation of different promoters[199-201] and to generating desmosines and isodesmosines in histone H1[89], which are a result of condensation of different lysine residues after LOX family members activity. In fact, previous work from our lab used the MNase assay to demonstrate that LOXL2 overexpression induces chromatin condensation at global levels (data not shown). These data reinforce the idea that histone H1 is a LOXL2 substrate; however, it would be interesting to investigate how LOX and LOXL2, which have very different protein sequences, could act on the same substrate. Both enzymes share a C-terminal domain, where the catalytic activity is conserved, but differ in their N-terminal region, which seems to confer specificity to protein-protein interactions. Hence, proteins involved in complexes with these enzymes may be responsible for providing the protein substrate specificity, with lysyl oxidases providing specificity for the substrate residue reaction. Thus, it is worth the effort to further analyse and characterize complexes containing different lysyl oxidases and their chromatin binding sites.

LOX activity in histone H1 has mostly been related to transcription activation, as lysine deamination produces a loss of positive charge, which would favour H1 detachment from the target DNA and consequent chromatin opening in target regions. In fact, indications of global chromatin decondensation have been shown in LOX overexpressing conditions[88]. However, effects at physiological levels of LOX enzymes on histones should be addressed and investigated in the complex context of chromatin structure and dynamics rather than based on the simplistic idea of histone H1 net charge. The fact that LOXL2 deaminated lysines can condensate with other deaminated and non-deaminated lysines suggests a new possible regulation mechanism for chromatin organization. As happens with other chromatin spatial organizations, such as Polycomb bodies, condensation of distant lysines in chromatin fibre may conform transcriptional regulation–associated domains or could change compaction status of chromatin due to condensation of subsequent H1 proteins. The phosphorylation status of H1 affects chromatin condensation, and this effect is related to structural changes rather than changes in net charge. For instance, hyperphosphorylation of H1 in metaphase chromatin induces condensation, and partial phosphorylation in interphase is involved in chromatin relaxation[202].

LOXL2 could also induce chromatin condensation through nucleosome assembly. In fact, previous data from our group showed that proteins involved in chromatin assembly often interact with LOXL2[91], suggesting a relation between this enzyme and nucleosome assembly. These LOXL2 interactors include p48 and p60, two of the three subunits of the CAF-1 chromatin assembly complex, which is responsible for histone H3 and H4 incorporation to chromatin[203]. Indeed, CAF-1 has been related to HP1α binding to pericentromeric regions[204]. Further, since CAF-1 interacts with HP1α and SUV39H1 (H3K9 methyltransferase) *in vivo*[205,206], it has been proposed that H3-H4 dimers or tetramers complexed with CAF-1 and HP1α may be methylated by SUV39H1 and deposited onto chromatin. HP1α could then be transferred to the methylated H3K9 residue and lead to formation and propagation of repressive chromatin. Nevertheless, how lysine condensation could be reversed is unclear. Histone exchange is a plausible mechanism for removing this modification. This exchange could be followed by a new unmodified or specifically modified histone or by a histone variant.

We also identified candidates related to DNA damage response (DDR), including PARP3, RAD50 and MRE11. These proteins are related to different DNA repair functions. For instance, RAD50 and MRE11 form a complex together with NBS1, called the MRN complex. This complex is a sensor of double-strand breaks and controls DNA damage

response through the activation of ATM kinase[207]. Interestingly, we found the SET7/9 recognition motif site in RAD50, indicating that this protein could be methylated and therefore could be a good candidate to be a LOXL2 substrate. While a lysine monomethylation for RAD50 has been recently described[208], its function is unclear; note that this is not the same lysine as the lysine candidate for SET7/9 methylation. It should be noted that the presence of this lysine modification does not exclude the possibility that 1) other lysines are also methylated in these proteins, or 2) the described lysine could alternatively be oxidized by LOXL2 or other oxidase enzymes.

This characterized methylation is localized in the N-terminal part of the protein, which suggests it is important in protein structure dynamics. The N-terminal and C-terminal regions of RAD50 fold back on themselves and bind to DNA to form the "RAD50 hook" domain, which mediates the formation of MRN complex assemblies[207]. PTMs of this region could affect DNA binding and complex formation. In general, lysine methylation in non-histone proteins seems to act in a similar way to that of histones, by modulating interactions with proteins containing "reader" domains. Methylation of non-histone proteins is likely to function by regulating protein-protein interactions in a similar way. There are several examples of lysine methylation inducing protein-protein binding, such as p53 with Tip60[209] and RB with HP1[210]. If RAD50 methylation follows the same

mechanism, it probably facilitates complex assembly. Oxidation of the residue could then impair this formation, allowing it to be regulated if unnecessary, or promote disassembly, thereby switching it off after it has achieved its purpose in the DNA damage response. It is likely that the rest of the complex co-immunoprecipitated with RAD50, although we cannot discard the possibility that it is indeed a direct substrate of LOXL2.

DDR is essential for maintaining the genomic integrity of the cell, and it is disrupted in many cancers[211]. This effect underlies the genomic instability that accompanies tumorigenesis and cancer progression. It has been proposed that DDR defects in tumour development and progression are positively selected by the need to tolerate oncogene-induced replication stress[212] or by the adaptive advantage provided by an increased mutation rate during tumour evolution[213]. However, in most cases, the specific underlying effects are poorly characterized[214,215]. In some cases, defects are not related to proteins directly involved in DDR but rather to its modulators. Along these lines, one could imagine that LOXL2 activity on this complex could affect its regulation in cancer, taking into account the relation between LOXL2 expression and tumour progression. In this scenario, inhibition of LOXL2 activity could be exploded therapeutically, as DDR defects are causative and permissive of disease. In fact, we have seen that neither knockdown of LOXL2 nor treatment with

LOXL2 inhibitor affects ESC survival unless cells are treated with RA; this is in stark contrast to treating the MDA-MB-231 cell line (derived from human breast adenocarcinomas) with LOXL2 inhibitor or depleting LOXL2 expression, both of which cause significant cell death. This effect could be related to alterations that this cancer cell line may have acquired in DDR, for which LOXL2 could play a role; indeed, these cells also have increased LOXL2 protein levels (data not shown). In this line, drugs targeting LOXL2 could be exploited therapeutically; in fact, there are already several drugs that target proteins to modulate DDR indirectly, such as DNMT1 and DNA polymerase[216].

Our mass spectrometry analyses also detected cytoplasmic proteins related to metabolism. Since our samples were enriched in nuclear fractions but did not completely exclude cytoplasmic fractions, it is not unexpected to find putative cytoplasmic substrates. However, since it was from a reduced fraction, we probably only detected the most abundant or easiest-to-detect cytoplasmic proteins. Hence, a different approach design would be required to identify properly cytoplasmic LOXL2 substrates. The candidates we detected are the GPI, TKT and TALDO enzymes, which are related to several pathways including pentose phosphate pathway, carbon metabolism and biosynthesis of amino acids. Moreover, TKT is one of the proteins for which we found a lysine suitable for SET7/9 methylation. However, no

methylations have been described for this enzyme. TKT is enzymatically active within the pentose phosphate pathway, which is an alternative metabolic pathway for glucose breakdown related to producing NADPH and pentoses used in nucleotide synthesis.

Because of the large biosynthetic demands of rapidly growing tumour cells and their need to adapt to stressful environments, the pentose phosphate pathway has been suggested to promote cancer progression and therapy resistance[217]. Many enzymes involved in this pathway are associated with malignancy[218] and one of them, G6PD, has been found to be negatively regulated by the tumour suppressor p53[219]. Taking into account the role of LOXL2 in cancer progression, it does not seem unreasonable to think that LOXL2 could be promoting metabolism reprogramming of cancer cells and allowing higher cell proliferation by inducing pentose phosphate pathway activity. Moreover, GPI also acts as an autocrine motility factor for cancer cells, stimulating metastasis when secreted. In fact, when the enzyme is overexpressed, it is able to induce EMT through upregulation of ZEB1 and ZEB2 genes expression[220]. How oxidation of these enzymes could be related to higher activity is not clear. One option would be that lysine oxidation abolishes ubiquitination and increases protein stability, leading to higher metabolic activity. In any case, the possibility that LOXL2

could have roles in such different and crucial cellular pathways deserves further research.

Lastly, we found TFIID complex as most confident candidate for LOXL2 oxidation, since it was a significantly enriched in the interactome module as well as enriched in the KEGG pathway analysis. TFIID is a general transcription factor responsible for RNAPII recruitment to gene promoters and thus related to active transcription. Hence, we approximated that LOXL2 oxidation could negatively regulate TFIID function to repress gene transcription since LOXL2 is a transcription repressor[91]. When investigating this possibility, we found that the TAF10 protein, one of the identified subunits of the TFIID complex, had already been described to have a lysine methylation transferred by SET7/9 methyltransferase[177], which is also able to methylate lysine 4 in histone H3. Since LOXL2 specifically recognizes trimethylated lysine 4 in histone H3, we decided to investigate possible TAF10 oxidation by LOXL2.

2. LOXL2 oxidizes the TAF10 transcription factor

After validating TAF10 oxidation detected by mass spectrometry, we demonstrated that LOXL2 interacted particularly with a highly posttranslationally modified TAF10. Mass spectrometry sequencing of modified TAF10 showed several lysine residues that had been mono-, di-, trimethylated or oxidized. However, we were not able to detect a unique candidate since there were more than one oxidized lysines, so that the specific lysine residue(s) oxidized by LOXL2 remain to be determined. TAF10 methylation is mainly restricted to the histone-fold domain (HFD), which undergoes a high degree of PTMs as we have experimentally demonstrated. Different lysines were found methylated and oxidized in this region, making it even more difficult to determine the specific substrate lysine residue. We know that mutation of K189 to glutamine abolishes TAF10 methylation levels completely and represses TAF10 methylation dependent-genes transcription[177], suggesting that K189 is required for the modification of the other lysines in the HFD. Thus, K189 methylation could produce a seeding effect that extends methylation to the other lysines in the domain. In this context, K189 could be a good candidate for LOXL2 oxidation. Interestingly, K189 is very close in sequence to other lysines that are also oxidized, suggesting the possibility that LOXL2 is able to oxidize all of them to completely abolish lysine methylation in the HFD, or that a first lysine oxidation

could have a seeding effect to induce recruitment of other oxidising enzymes to target the other lysine residues.

We also showed that LOXL2 catalyses oxidation of methylated TAF10 *in vitro* and *in vivo*. LOXL2 showed similar activity towards both wild-type and mutant unmodified TAF10 *in vitro* (not shown); however, this apparent lack of specificity may be due to the excess of recombinant LOXL2 enzyme present with respect to the amount of TAF10 used as substrate, which was immunoprecipitated from nuclear fractions. In fact, although LOXL2 showed similar binding to both wild-type and mutant TAF10 *in vivo*, LOXL2 oxidized the modified form of wild-type TAF10 to a much higher extent than the non-methylated mutant. This results suggest that lysine methylation in TAF10 gives the specificity for oxidation but not interaction.

Although further studies are required to validate and elucidate the role of TAF10 posttranslational modifications, a recent paper shows that HFD is key in promoting the TAF10-TAF8 interaction[77], which in turn triggers the transition from a symmetric to an asymmetric TFIID complex. Since TAF10 is necessary for this transition and to maintain the TFIID integrity[79,181], and since HFD of different TFIID subunits mediates many of the interactions within TFIID[71], we proposed that TAF10 oxidation may affect the final TFIID composition. In this way, the formation of an aldehyde group

after the LOXL2 reaction could modify the macromolecular structure of TAF10 and the interactions of the oxidized protein. However, we did not observe significant changes in the general TFIID composition in HEK293T cells infected with either wild-type or mutant LOXL2. Oxidized TAF10 might correspond to a fraction of the total TAF10 (e.g., methylated TAF10, known to specifically regulate a subset of genes[177]), which would hinder our ability to detect changes in the composition of this general complex. In fact, we were not able to detect TBP, TAF11 or TAF13 in the co-immunoprecipitated sample when oxidized TAF10 was pulled down, suggesting that this modification compromises the interaction with some members of the complex. TAF10 HFD oxidation may also affect its cellular localization, since it has been described that TAF10, which lacks a nuclear localization signal, depends entirely on interaction with other subunits through HFD for its nuclear transport. Indeed, the TAF10 protein is not retained in the nucleus if it is not interacting with other subunits[180]. Hence, TAF10 incorporation into the TFIID complex might be blocked by preventing an interaction necessary for nuclear retention/transport through an oxidized HFD residue. Cellular subfractionation assays should be performed to clarify this point and to better elucidate how PTMs can regulate TAF10 localization dynamics.

3. LOXL2 represses TFIID dependent genes through TAF10 oxidation

We propose that oxidation of the methylated TAF10 contained in a TFIID complex inhibits the ability of TFIID to activate transcription of its regulated genes, since its TAF10 component is no longer maintained in the complex. Our results show that only described TAF10 methylation-dependent genes are affected by LOXL2 activity, and that oxidized TAF10 is enriched in promoters when these are inactivated and just before TAF10 is released. Whether the TFIID complexes that have lost their methylated TAF10 go on to form an incomplete TFIID complex, or whether the subunits are redistributed into other TFIID complexes, is not known. Finally, not only TAF10, but also three other members of the complex, namely, TBP, TAF1 and TAF11, as well as RNAPII were released from the TFIID-sensitive genes, suggesting that the entire TFIID complex detached from those promoters, thereby impairing pre-initiation complex formation.

The release of TAF10, TBP, TAF1 and TAF11 may also promote the loss of the SAGA complex, another large coactivator in which TAF10 and TBP are very relevant components[221,222]. This release would be accompanied by a reduction in the acetylated levels of histones, facilitating loss of the preinitiation complex and gene repression. Whether TAF10 oxidation can also affect the TBP-free TAF complex

(TFTC) is not known. Since this complex is related to GCN5 histone acetyltransferase, that would imply that LOXL2 activity would led to a reduction of acetylated histones, and thus to gene transcription repression. Since the presence of oxidized histone H3 is also increased in these promoters, as we showed, it is likely that LOXL2 plays an additional role in catalysing histone H3K4me3 oxidation and thereby in antagonizing methylation of this residue[91]. Furthermore, H3K4me3 provides a high-affinity binding site for TFIID through TAF3 subunit[74]; therefore, oxidation of histone H3 would also impair TFIID binding.

Since LOXL2 activity represses TAF10 methylation–dependent genes, the effects may vary between cell lines and different physiological contexts depending on levels and binding sites of methylated TAF10. For example, in HEK293T cells, overexpression of LOXL2 leads to repression of TAF10 methylation–dependent genes but has no effect on cell survival, proliferation or identity. However, in mESCs, LOXL2 overexpression leads to morphological changes due to pluripotent state loss. In this latter case, neither TAF10 methylation nor methylation-dependent genes have already been described. We have shown that TAF10 is methylated in ESCs, and that it is oxidized when LOXL2 is overexpressed. Moreover, pluripotency gene expression, which is dependent on TFIID[80], is repressed in wild-type LOXL2 but not in mutant LOXL2–expressing cells. Although we do not know if they

also depend on TAF10 methylation to be transcribed, we demonstrated that LOXL2 is present in these promoters, inducing TAF10 oxidation and TFIID release from them. Because LOXL2 has a preference for oxidizing the methylated form of TAF10 *in vivo*, it may be that the TAF10 present in pluripotency gene promoters, which is oxidized by LOXL2 and thereby released, may also be methylated.

Furthermore, we reasoned that LOXL2 activity affects TAF10 binding to TFIID, and that this may affect levels of functional TFIID complex by disrupting the pool of methylated TAF10–containing TFIID. Even if all TFIID subunits were maintained at similar levels, a lack of TAF10 would impair the transition from a core TFIID to the complete and functional holo-TFIID complex, since recruitment of the required subunits would be impaired. Conditions in which the holo-TFIID complex levels are reduced were found to also affect pluripotency gene transcription, since it has been described that a decrease in global TFIID protein levels represses their expression[80]. This study proposed that pluripotency gene promoters have a low affinity for TFIID, which makes them more sensitive to a reduction of global levels of the complex[80]; this situation could also occur under our conditions due to a reduction of TAF10 levels. Because TAF10 methylation is described to increase RNAPII affinity for TFIID complex[177], a possibility is that TFIID containing methylated TAF10 may be present in these promoters that have a lower affinity for TFIID, which would

help increase pre-initiation complex formation under these disadvantageous conditions. In this context, target promoters would be more sensitive to LOXL2 activity, both because LOXL2 prefers to oxidize methylated TAF10 and because LOXL2 could decrease total holo-TFIID levels by inducing TAF10 degradation. It should be noted that (as explained above) loss of H3K4me3 would also impair new TFIID binding to target promoters, leading to a strong gene repression by LOXL2 enzyme. It now remains to be determined in which cells, and to which extent, TAF10 methylation and oxidation are crucial.

4. LOXL2 activity is required in ESC differentiation

We demonstrated that RA treatment leads to increased LOXL2 levels concomitant with increased TAF10 oxidation levels and repression of pluripotency genes. We determined that TAF10 oxidation precedes a reduction in TAF10 global levels and demonstrated that this is due to protein degradation, since no changes were observed at transcriptional level, and it was abolished in the presence of a proteasome inhibitor. Moreover, we showed that this degradation requires LOXL2, since it was recovered in LOXL2-depleted ESCs treated with RA. How TAF10 oxidation induces protein degradation is still unknown. A possible explanation is that TAF10 oxidation prevents its interaction with other subunits. As mentioned above, the HFD is key to TAF10's interaction with other TFIID subunits, and we have shown that when this domain is oxidized, this interaction is lost. A non-interacting TAF10 may be a better target for ubiquitin ligases, leading to its degradation via the proteasome. In addition, it is also possible that TAF10 ubiquitin ligases localize in the cytoplasm, and that forced TAF10 transport to this compartment due to a lack of interactors makes it sensitive to degradation.

Furthermore, why TAF10 degradation occurs upon ESC differentiation is also an open question. It has been demonstrated that a lack of TAF10 leads to cell cycle arrest

and cell death by apoptosis in mouse F9 embryonic carcinoma cells[223], and that TAF10 is required for early mouse development and survival of the pluripotent inner cell mass but not for survival of mouse trophoblast cells[181]. Moreover, in a mouse model with a liver-specific disruption of TAF10, inactivation of TAF10 in hepatocytes resulted in the dissociation of TFIID into individual components, which correlated with the downregulation of most hepatocyte-specific genes during embryonic life and a defect in liver organogenesis. However, the transcription of less than 5% of active genes was affected by TAF10 inactivation and TFIID disassembly in adult liver[79]. Together, these data suggest that TAF10 is required for transcription of a subset of genes essential for stem cells, but once cells have differentiated, the number of genes affected by TAF10 depletion is low and that these do not affect cell viability. Hence, high levels of TAF10 protein may be required for cell viability and pluripotent state of stem cells, but they are not necessary in differentiated cells and protein levels decrease upon differentiation.

We have demonstrated that LOXL2 represses pluripotency genes, and that LOXL2 deficiency affects the transcription of pluripotency genes that remain expressed in the presence of RA. Moreover, LOXL2-depleted ESCs showed lower viability only when induced to differentiate with RA treatment. After several days of treatment, LOXL2-depleted cells that were able to differentiate were shown to have recovered LOXL2

expression. These results suggest that LOXL2 is required for a proper RA-induced ESC differentiation. However, it is not clear why the viability of LOXL2-deficient ESCs decreases when incubated with RA. Since pluripotency genes have not yet been properly downregulated, it is possible that these cells detect two activated gene expression patterns: the one that signals for pluripotency maintenance, and the one that signals for differentiation. In this situation, having more than one pathway activated, with each trying to impose its transcriptome expression pattern to define cell identity, may cause cells to activate checkpoints that lead to cell death.

We performed a transcriptome analysis to identify genes that are silenced upon RA treatment and depend on LOXL2 for this repression. We found out that a striking 40% of genes changing upon RA treatment depended on LOXL2 presence to do so. This result highlights the importance of LOXL2 in this differentiation pathway and is in accordance with the critical role for LOXL2 in survival in RA-induced differentiation that we showed. As expected, affected genes were related to embryonic development and nervous system development.

We also identified TFIID-dependent genes in our transcriptome that are downregulated during RA differentiation in a LOXL2-dependent manner. This subset of genes represented 20% of total TFIID-dependent genes downregulated upon RA treatment. Whether these genes

require a methylated TAF10 for transcription is not known, and an antibody for TAF10 methylation should be generated to answer this question. However, the fact that they are highly sensitive to TAF5 knockdown suggests that, even if they do not require TAF10 methylation for their transcription, they respond to total levels of TFIID and thus are likely to be affected by the pronounced reduction of TAF10 protein levels in RA differentiating conditions.

Although LOXL2 is transcriptionally induced during differentiation, it is also present in basal conditions in ESCs as well as in the pluripotency gene promoters. Transcriptome profile in LOXL2 knockdown ESCs showed numerous developmental genes affected in basal conditions related to the development of the cardiovascular system, visual system, organism and reproductive system. Moreover, we demonstrated upregulation of pluripotency factors when LOXL2 was downregulated in basal conditions in ESCs. These results suggest a role for LOXL2 in regulating pluripotency network dynamics, and probably in maintaining the pluripotent state, when present at low levels in ESCs.

Under differentiation conditions, LOXL2 is expressed and located in the promoters, but we only detected oxidation of histone H3 after RA treatment, which suggests that LOXL2 needs RA to become fully active in the promoters we checked; whether this activation depends on other

coactivators or is due to increase in LOXL2 levels is still unknown. Since the LOXL2 gene promoter lacks retinoic acid response elements, it is unlikely that LOXL2 is an early RA response gene, so that upregulation of transcriptional activity must be an indirect regulation by RA early response genes, such as HoxA1 or other Hox genes[170].

Further research is needed to verify if LOXL2 could present a dosage effect. At basal levels in ESCs, both pluripotency factors and development related factors would be repressed by LOXL2. Upon RA signalling, higher LOXL2 levels, together with possible higher enzymatic activity, would highly repress pluripotency genes and ground state would no longer be maintained. Developmental genes repressed by LOXL2 at basal levels, which are not related to neural commitment, could also be more strongly repressed to ensure a correct differentiation into neural development and avoiding other differentiation pathways.

It is also worth considering whether LOXL2 has a regulatory role in the RA signalling pathway. We have shown LOXL2 substrate candidates in the PPAR signalling pathway, where nuclear hormone receptors interact with retinoic acid receptors to regulate gene transcription. Moreover, a lysine trimethylation for retinoic acid receptor alpha (RARα) has been described, which induces binding of RARα to its coactivators[224]. As LOXL2 has a preference to oxidize

trimethylated lysines, and as we have localized its activity in a pathway for gene expression pattern regulation by nuclear hormone receptors, it is reasonable to think that LOXL2 could oxidize RARα to regulate RA activity through a negative loop. In fact, LOXL2 activity is related to adult stem cells or multipotent cells in other differentiation pathways, while its expression is lost at terminal differentiation. As continued RA expression leads to a final differentiation to neurons, LOXL2 expression would be expected to decrease along these differentiation stages if it follows same pattern as other differentiation pathways[113,115]. In this context, LOXL2 may act negatively upon RA signalling activity to maintain multipotent stage of neural progenitor cells and impair terminal differentiation to neurons.

It is worth pointing out that these expression patterns correlate with LOXL2 expression patterns in cancer and its role in tumour growth, since tumour cells responsible for growth are commonly related to adult progenitor cells. However, even though LOXL2 is expressed in adult stem cells or progenitor cells, and its downregulation is required for proper differentiation, whether or not LOXL2 expression is required for progenitor cell induction or maintenance and the molecular mechanisms related to this process are still unknown.

5. LOXL2 has a role in embryonic development

Finally, we have demonstrated that Sox2 levels were abnormally upregulated in the absence of LOXL2 during zebrafish development, which affected normal neural differentiation, as observed by the morphological defects of the brain and the low survival rate of the embryos. This is not the first time that lysyl oxidase activity has been studied in zebrafish development. Eight Loxl proteins were identified in zebrafish so far, including orthologs for each mammalian LOXL2 protein, except LOXL4, and two additional Loxl proteins, Loxl5a and Loxl5b, which are closely related to Lox and Loxl1. The *lox* gene is expressed from 3 hpf to 5 dpf, clearly indicating a role in early embryonic development. In fact, at 48 hpf, it is expressed in the developing nervous system, eyes, pectoral fin and muscle. However, in contrast to *loxl2* knockdown, *lox* downregulation has different consequences as those observed in our model, including smaller heads[225]. In this case, although is clear that *lox* has a role in nervous system development, it is different to the one we showed for *loxl2a* and *loxl2b*. Moreover, knockdown of *loxl1* and *loxl5b* results in notochord distortion[185], indicating that other members of the family have a function more related to that of *lox*. This concurs with differentiation pathways studied in cell culture models, in which LOXL2 expression has different patterns, and sometimes even opposite patterns, to the ones of the other family members[113,115].

Although methylation and oxidation of TAF10 have not been demonstrated in this model, it is worth to notice that all lysine residues for which we characterized methylation and oxidation in human cells are conserved in zebrafish. This evidence highlights the importance of HFD in TAF10 regulation.

6. Lysine oxidation as a posttranslational modification

Dynamic posttranslational modifications serve to regulate protein-protein interactions, protein stability, protein localization and enzymatic activities. Protein methylation of lysine and arginine residues on non-histone proteins has emerged in the past few years as a prevalent posttranslational modification. In a simplified way, the process of dynamic methylation may be represented by a "writer-reader-eraser"-model[226]. In this context, methyltransferases would be writers, proteins containing methyl-binding domains would be readers and enzymes removing methyl marks including demethylases or deaminases, such as LOXL2, would be erasers. These dynamics lead to changes in the activity, localization or interactors of target proteins. An example is methylation of p53 by the SETD8 writer, which promotes binding of the reader 53BP1 through its tudor domains. This interaction can then be impaired by demethylation of p53 by the KDM1A eraser.

In a similar way, LOXL2 could be acting as an eraser, removing methylation from TAF10 protein. Although oxidation has been mainly considered to be an uncontrolled and non-enzymatic modification, many enzymatic oxidations occur on enzyme-selected lysine residues. In fact, histone demethylases have been described to remove methyl groups

through oxidative reactions in target residues and are referred to as amino oxidases. LSD1 itself, first described as a histone demethylase, is a FAD-dependent amino oxidase able to remove mono- and dimethylation from lysines that produces H_2O_2 as a subproduct of oxidative reaction[227], in a similar way to LOXL2. The JmjC domain-containing histone demethylases mediate oxidative demethylation by a radical attack involving Fe(II)[228]. Furthermore, in some cases, identifying demethylase candidates has focused on looking for proteins with domains capable of oxidative reactions[229]. However, unlike demethylases, which leave target residues unmodified and suitable to receive new methyl groups, LOXL2 leaves a modified lysine as a product in substrate proteins, an allysine. This activity resembles that from the protein arginine deiminase (PAD) family of proteins, which are able to perform arginine deimination and leave a modified residue called citrulline as a product. Indeed, the PAD family also contains histone-modifying enzymes, since PAD4 has been described to erase arginine methylation from histone H3 through deimination[230,231]. Interestingly, PAD4 also shows activity for other histones[232] as well as for non-histone substrates[233].

In our model, LOXL2 can be also considered as a writer of allysine residue, and TAF10 oxidation as a novel controlled protein modification that demonstrates a critical role for protein oxidation in regulation of protein function. We have

demonstrated that LOXL2 activity induces TAF10 degradation; thus, one of the outcomes of this oxidation could be signalling to protein degradation. Nevertheless, resulting functions of this posttranslational modification deserves further study. It seems likely that there are other reader proteins able to recognize oxidized lysines yet to be discovered, whose characterization would shed light on the biological roles of lysine oxidation.

CONCLUDING REMARKS

Protein function is often regulated and controlled by posttranslational modifications, such as oxidation. Although oxidation has been mainly considered to be uncontrolled and non-enzymatic, many enzymatic oxidations occur on enzyme-selected lysine residues. This is the case for the LOXL2 enzyme, which is able to oxidize lysines by converting the ε-amino groups into aldehyde groups.

Following the initial objectives of this thesis, our work has generated a list of putative nuclear and cytoplasmic LOXL2 substrates. We have demonstrated that one of those candidates, the TAF10 transcription factor member of the TFIID general transcription, interacts with the LOXL2 enzyme. Interacting TAF10 protein shows a high degree of posttranslational modifications, including lysine methylation and lysine oxidation, all of which have been described here for the first with the exception of lysine monomethylation in K189. Moreover, our findings indicate that LOXL2 directly oxidizes TAF10 protein preferentially in its modified form, affecting the transcriptional activity of TAF10 methylation–dependent genes by releasing the TFIID complex from target promoters (*Figure CR.1*).

This repression mechanism has proven to be key in pluripotency maintenance in ESCs, in which pluripotency gene transcription depends on TFIID complex activity. LOXL2 expression reduces the pluripotent state of ESCs as a

consequence of pluripotency gene repression. Under RA-induced differentiation conditions, ESCs upregulate the LOXL2 protein, leading to TAF10 oxidation and degradation. Concomitantly, TFIID is released from pluripotency gene promoters, resulting in transcription repression. Furthermore, we have demonstrated that in LOXL2 depleted conditions, ESCs do not correctly repress pluripotency genes upon RA treatment. Global transcriptome analysis has also shown additional subset of genes that are repressed upon differentiation in a LOXL2-dependent manner. All these data reinforce the idea that LOXL2 plays a role in repressing key genes for ESC differentiation. This has been further confirmed by analysing how LOXL2 depletion affects zebrafish embryo development: it leads to an overexpression of the Oct4 pluripotency factor and a failed neural commitment development, as evidenced by loss of the NeuroD marker.

In conclusion, we have shown that lysine oxidation of the transcription factor TAF10 by LOXL2 is a controlled protein modification, and that the deregulation of this process disturbs the balance between pluripotency and differentiation. However, further studies will be required to elucidate the biological relevance of lysine oxidation as a posttranslational modification, not only in LOXL2 substrates but also in other proteins subject to this modification, in order to answer still unclear key questions: what is the molecular effect of lysine

oxidation? To which extent are proteins oxidized? Are there specific domains for allysine recognition? Are oxidation levels altered in pathological conditions, like cancer, as observed for other modifications, such as methylation? Answering these questions will help us to better understand the dynamics of protein posttranslational modifications, the resulting protein function control and the consequences of misregulation.

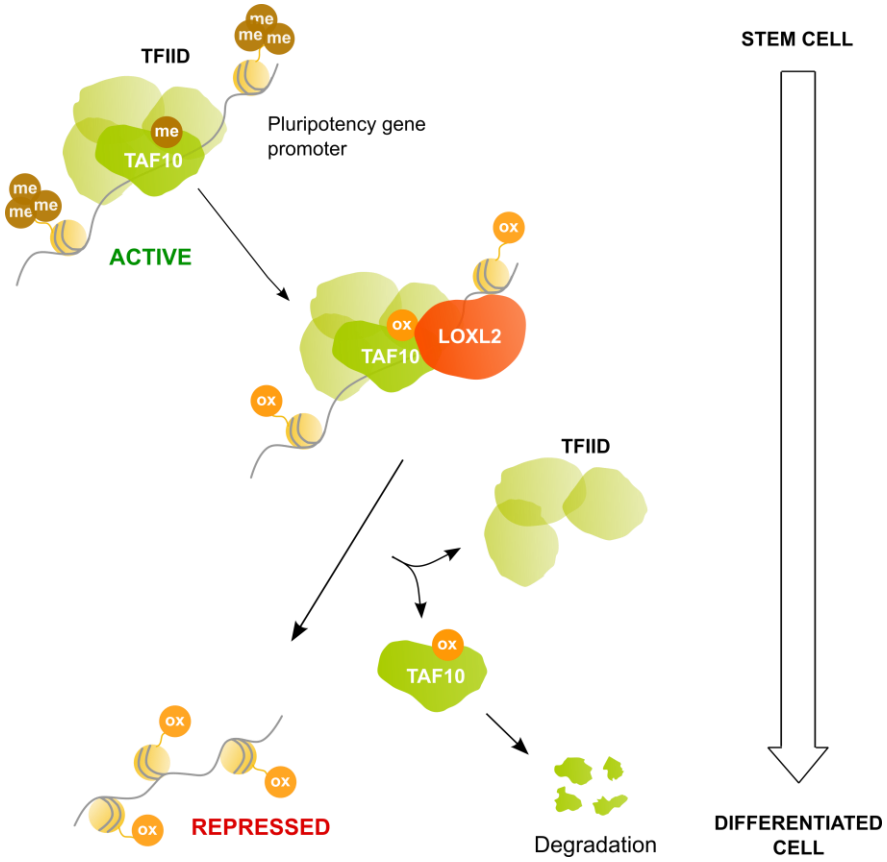

Figure CR.1. Representation of working model. LOXL2 oxidation of TAF10 induces its release from target promoters, leading to a block in TFIID-dependent gene transcription reinforced by histone H3 oxidation. Released oxidized TAF10 can no longer interact with other subunits of the complex and is degraded. In ESCs, this results in the inactivation of the pluripotency genes and loss of the pluripotent capacity, which occurs physiologically when ESCs take the RA differentiation pathway.

MATERIALS & METHODS

1. Cell culture

HEK293T cells and HEK293 gag-pol cells were maintained in Dulbecco's modified Eagle's medium (Invitrogen) with 10% fetal bovine serum (Invitrogen) at 37ºC in 5% CO2. E14Tg2A ESCs were cultured feeder–free in plates coated with 0.1% of gelatin in Dubelcco's Modified

Eagle Medium supplemented with **β**-mercaptoethanol, sodium pyruvate, essential amino acids, glutaMAX, 15% fetal bovine serum (Hyclone) and leukaemia inhibitory factor (LIF) at 37ºC in 5% CO2.

1.1. ESC differentiation

For ESC differentiation to neural progenitors, cells were cultured with 1 □M of *all-trans* retinoic acid (ATRA) (Sigma) and no LIF for 1, 2, 4 and 6 days, with the media replaced every 24 h.

2. Cell transfection and infection

2.1. Transfection

HEK293T cells were seeded in p100 plates with 10 ml medium and grown for 24 hr until 70% confluence was reached. Cells were transfected with 10 μg of pcDNA3-hLOXL2Flag or pXJ41-hTAF10HA vectors using polyethylenimine polymer (PEI; Polysciences Inc). A mixture of 1.5 ml 150 mM NaCl, DNA and 78 μl PEI was incubated for 15 min at room temperature and then added drop-wise to cell plates.

2.2. Retrovirus production and infection

For retroviral infections, HEK293 gag-pol cells were used. Cells were seeded in p150 plates with 14 ml medium, grown to 70% confluence and then transfected (day 0) by adding drop-wise, a mixture of 3.3 ml 150 mM NaCl, DNA (2.5 μg of pCMV-VSV-G and 7.5 μg of pMSCV, pMSCV-LOXL2 wt-FLAG or pMSCV-LOXL2 mut-FLAG ires GFP vectors) and 166 μl PEI (Polysciences Inc) that were preincubated for 15 min at room temperature. The transfection medium was replaced with 12 ml fresh medium after 24 hr (day 1). The cell-conditioned medium at day 2 and 3 was filtered through 45 μm membrane filters (Millipore) and collected together. For

virus concentration, 8 ml of retroviral concentrator (#631456; Clontech) was added and the mixture was incubated overnight at 4ºC. Precipitated virus was centrifuged 45 min at 3,000 rpm and pellet was resuspended in 1 ml of fresh medium. 100 μl aliquot of virus was used to infect target cells together with 8 μg/mL polybrene. 24 hr after infection, medium of the target cells was changed and puromycin added to select infected cells for 72 hr (1μg/ml for HEK293T cells and 2μg/ml for mESCs).

2.3. Lentivirus production and infection

For lentiviral infection, HEK293T cells were used to produce viral particles. Cells were transfected as described for HEK293 gag-pol cells with a DNA mixture comprising 50% pLKO-shControl/shLoxl2, 10% pCMV-VSVG, 30% pMDLg/pRRE and 10% pRSV rev. The rest of the procedure was the same as for retroviral infection of cells but using a lentiviral concentrator (#631232; Clontech). In order to silence Loxl2, the following shControl and shLoxl2 sequences (Sigma) were used: 5′-CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT-3′ and 5′-CCGGCCAAATAGAGAGCCTAAATATCTCGAGATATTTAGGCTCTCTATTTGGTTTTTG-3′ respectively.

3. Protein analysis

Before analysis of protein extracts, quantification in triplicates using the *DC* Protein Array kit (Lowry method; Bio-Rad) was carried out. Prior to addition of the desired buffer for lysis, cells were washed three times with cold PBS and scraped in the plate with the buffer.

3.1. Total extracts

Total cell extracts were obtained with 2% SDS lysis buffer. Samples were kept at room temperature to avoid precipitation of the SDS, syringed five times, centrifuged at 13,200 rpm for 10 min and boiled at 95ºC for 3 min.

2% SDS lysis buffer
2% SDS
50 mM Tris-HCl pH 7.5
10% glycerol

3.2. Nuclear extracts

Nuclear extracts were obtained lysing cells with Soft-lysis buffer and kept 5 min on ice. Then, samples were centrifuged at 3,000 rpm for 15 min and the supernatant was discarded.

The pellet containing nuclear fraction was lysed in High salt-lysis buffer for 30 min at 4ºC, and samples were centrifuged at 13,000 rpm for 10 min.

Soft-lysis buffer	**High salt-lysis buffer**
50 mM Tris	30 mM HEPES pH 7.4
2 mM EDTA	350 mM NaCl
0.1% NP-40	1 mM $MgCl_2$
10% Glycerol	0.5% Triton X-100
Protease and phosphatase inhibitors	10% Glycerol
	Protease and phosphatase inhibitors

3.3. Western blot

Protein was analysed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) by loading the samples previously mixed with 5x Loading buffer and boiled at 95ºC for 3 min. Gels had a 7.5-15% polyacrylamide concentration. The Mini-Protean System (Bio-Rad) was used to run gels in TGS buffer that were then transferred to Protran nitrocellulose membranes (Whatman) during 60-90 min depending on the molecular weight of the protein using Transfer buffer.

5x Loading buffer

250 mM Tris-HCl pH 6.8
10% SDS
0.02% Bromophenol blue
50% Glycerol
20% β-mercaptoethanol

Transfer buffer

50 mM Tris-OH
386 mM Glycine
0.1% SDS
20% methanol

TGS buffer

25 mM Tris-OH pH 8.3
192 mM Glycine
5% SDS

TBS-T

25 mM Tris-HCl pH 7.5
137 mM NaCl
0.1% Tween-20

Ponceau S stain

0.5% Ponceau
1% acetic acid

Prior to blocking, and to ensure that protein was loaded and transferred correctly onto the membrane, Ponceau S staining was performed to each membrane. Membranes were placed directly from the Transfer buffer to Ponceau S staining and rocked for 5 min. The solution was removed and various washes were performed to remove excess stain with distilled water.

Membranes were blocked in 5% skimmed milk in TBS-T for 1 hr and incubated in the desired antibody overnight at 4ºC. After three 10 min washes with TBS-T, Horseradish peroxidase (HRP)-combined secondary antibody (Dako) was diluted in 5% skimmed milk and the membrane was incubated

for 1 hr at room temperature. Prior to developing three more 10 min washes with TBS-T were performed. Membranes were developed using Luminatan Western HRP Substrates (Millipore) and exposed on AgfaCurix or Hyperfilms ECL (Amersham) for proteins that were more difficult to detect.

3.4. Pull-down assays for oxidized proteins detection

Cells were cotransfected with pcDNA3-hLOXL2wt-Flag and pXJ41hTAF10wt/K189Q-HA in case of HEK293T cells and infected with pMSCV-hLOXL2wt/mut or treated with RA in case of ES cells. 24 hr after transfection of HEK293T and 72 hr after infection and selection of ES cells or RA treatment at indicated time points, nuclear extracts were obtained as described above. Hydrazide-biotin was then added to the samples at 10 mM final concentration and incubated at 25ºC for 2 hours[91]. Streptavidin-magnetic beads (New England Biolabs) were added to pull down biotinilated proteins for 1 hr at 4ºC. After washing once with 150 mM Salt washing buffer and twice with 300 mM Salt washing buffer, oxidized proteins were eluted with Loading buffer 2x and boiling for 3 min at 95ºC.

150 mM washing buffer	**300 mM washing buffer**
20 mM HEPES pH 7.4	20 mM HEPES pH 7.4
150 mM NaCl	300 mM NaCl
1 mM $MgCl_2$	1 mM $MgCl_2$
0.5% Triton X-100	0.5% Triton X-100
10% Glycerol	10% Glycerol

3.5. Immunoprecipitation (IP)

3.5.1. IP of exogenous and endogenous proteins

Immunoprecipitation was carried out using HEK293T cells seeded in p100 plates. In case of exogenous proteins, cells were grown at 70% confluence they were transfected using PEI with 4 μg of the indicated plasmids. 24 hr after transfection cells were washed with cold PBS and lysed with 500 μl Soft-lysis buffer. The lysates were incubated at 4ºC for 5 min and centrifuged at 3,000 rpm and 4ºC for 15 min. The supernatant was discarded and the pellet was lysed with 300 μl High salt-lysis buffer followed by 30 min incubation at 4ºC. Lysates were then centrifuged for 10 min at 13,200 rpm. Supernatant NaCl concentration was reduced to 300 mM NaCl with Balance buffer. The lysate was subjected to immunoprecipitation with α-Flag M2 agarose beads (20 μl/sample; Sigma) for 4 hr in case of Flag immunoprecipitation or with the indicated antibody overnight

at 4ºC followed by collection of immunocomplexes with protein A or G-magnetic beads (Millipore), depending on primary antibody species, with 1 hr rotation at 4ºC. Samples were washed once with 150 mM Salt washing buffer and twice with 300 mM Salt washing buffer. Precipitated complexes were then eluted with Loading buffer 2x.

Balance buffer

20 mM HEPES pH 7.4

1 mM $MgCl_2$

10 mM KCl

3.5.2. RelP for TAF10 trimethylation and oxidation detection

Immunoprecipitation was carried out using HEK293T and ES cells seeded in p100 plates. In case of exogenous proteins, HEK293T cells were grown at 70% confluence they were transfected using PEI with 4 μg of the indicated plasmids. 24 hr after transfection cells were washed with cold PBS and lysed with 500 μl Soft-lysis buffer. The lysates were incubated at 4ºC for 5 min and centrifuged at 3,000 rpm and 4ºC for 15 min. The supernatant was discarded and the pellet was lysed with 300 μl High salt-lysis buffer followed by 30 min incubation at 4ºC. Lysates were then centrifuged for 10 min at 13,200 rpm. Supernatant NaCl concentration was reduced to

300 mM NaCl with Balance buffer. In case of oxTAF10 detection, at this point samples were incubated with hydrazyde-biotin for 2 hr at 25ºC.

The lysates were subjected to immunoprecipitation with α-HA (for exogenous TAF10) or α-TAF10 (for endogenous TAF10) antibodies for 2 hr in case of α-HA and overnight for α-TAF10 at 4ºC. Immunocomplexes were then collected with protein A or G-magnetic beads (Millipore), depending on primary antibody species, with 1 hr rotation at 4ºC. Samples were washed once with 150 mM Salt washing buffer and twice with 300 mM Salt washing buffer. Precipitated fractions were eluted in SDS 1% (diluted from SDS 10% in Phosphate buffer) for 1 hr at 37ºC. SDS concentration was then diluted at 0.1% with Phosphate buffer and incubated either with streptavidin-magnetic beads for 1 hr in case of TAF10 oxidation detection or with anti-N-ε-trimethyl lysine antibody agarose conjugates (30 μl/sample; Immunochem) overnight in case of TAF10 trimethylation detection at 4ºC. Samples were washed once with 150 mM Salt washing buffer and twice with 300 mM Salt washing buffer. Precipitated complexes were then eluted with Loading buffer 2x.

Phosphate buffer

0.1M Na_2HPO_4

0.1M NaH_2PO_4

4. Recombinant protein purification

pFastBac1 baculovirus vectors coding for LOXL2 wt-FLAG and LOXL2 mut-FLAG were used to generate recombinant bacmids using the Bac-to-Bac® system (Invitrogen). Isolated bacmids were transfected in Sf9 insect cells using Cell Fectin II reagent (Invitrogen) to generate high titer baculoviruses.

LOXL2-FLAG recombinant proteins (wild-type and mutant) were purified from Sf9 insect cells 2-3 days after infection. Cell pellet was resuspended in High salt lysis buffer and incubated in ice for 30 min. lysates were then centrifuged at 13,000 rpm and 4ºC for 10 min. Supernatant was then balanced to 200 mM NaCl with Balance buffer. Samples were incubated with FLAG M2 beads (1.5 ml beads for each 4 ml sample) for 4h at 4ºC and washed three times with 300 mM salt washing buffer and twice with 100 mM salt washing buffer. Elution was carried out with 300 μl of Elution buffer containing 1 μg/μl of FLAG peptide for 1 hr at 4ºC. A second elution round was performed adding 300 μl more of Elution buffer overnight at 4ºC.

High salt lysis buffer	**300 mM washing buffer**
20 mM HEPES pH 7.4	20 mM HEPES pH 7.4
10% Glycerol	10% Glycerol
350 mM NaCl	300 mM NaCl
1 mM $MgCl_2$	1 mM $MgCl_2$

0.5% Triton X-100
1 mM DTT
Protease and phosphatase inhibitors

100 mM salt washing buffer
20 mM HEPES pH 7.4
10% Glycerol
100 mM NaCl
1 mM $MgCl_2$
0.5% Triton X-100
1 mM DTT
Protease and phosphatase Inhibitors

0.5% Triton X-100
1 mM DTT
Protease and phosphatase inhibitors

Elution buffer
20 mM HEPES pH 7.4
100 mM NaCl
1 mM $MgCl_2$
1 mM DTT
Protease and phosphatase inhibitors
FLAG peptide (1 μg/μl)

5. *In vitro* reaction assays

In vitro experiments and further detection of oxidized and trimethylated TAF10 levels. First, TAF10 was isolated from HEK293T cells transfected with pXJ41-hTAF10-HA by HA immunoprecipitation but without the elution of precipitated proteins (i.e., complexes were still bound to the magnetic beads). TAF10 complex was then incubated with rLOXL2 wt or rLOXL2 mut in Phosphate buffer for 2 hr at 37ºC and 110 rpm. After washing three times with Phosphate buffer, TAF10 complexes were eluted in 1% SDS for 1 hr at 37ºC, and SDS was then diluted to 0.1% with Phosphate buffer. To detected oxidized TAF10, hydrazide-biotin was added at a 10mM final concentration, and the samples were incubated at 25ºC for 2 hr. Streptavidin-magnetic beads were added (20 μl/sample) for 30 min at 4ºC to pull down biotinylated proteins. After washing with phosphate buffer, precipitated complexes were eluted with 2x Loading buffer.

For trimethylated TAF10 detection, TAF10 was isolated from HEK293T cells transfected with pXJ41-hTAF10-HA by HA immunoprecipitation, but without the elution of precipitated proteins to keep the complex bound to magnetic beads. TAF10 complex was then incubated with rLOXL2 wt or rLOXL2 mut in Phosphate buffer for 2 hr at 37ºC and 110 rpm. After washing three times with Phosphate buffer, TAF10 complexes were eluted in 1% SDS for 1 hr at 37ºC. SDS was

diluted to 0.1% with Phosphate buffer, and anti-N-ε-trimethyl lysine antibody agarose conjugates (30 μl/sample; Immunochem) were added and incubated overnight at 4ºC. Samples were then washed twice with 300 mM Salt washing buffer and eluted with 2x Loading buffer.

6. Proteasome and LOXL2 inhibition assays

ES cells were treated with 1 mM ATRA as described above during the indicated time points. For proteasome inhibition, MG132 (Z-Leu-Leu-Leu-al) (Sigma) was added to 50 mM final concentration for 6 hr preceding cell collection. For LOXL2 inhibition, zonisamide sodium salt (Sigma) was added to 5 μM, 10 μM and 50 μM final concentrations for the same time as ATRA treatment, 1 day. Total extracts were obtained with 2% SDS lysis buffer, and the resulting proteins were resolved by SDS-polyacrylamide gel electrophoresis and analyzed by western blot with indicated antibodies.

7. RNA analysis

7.1. Phenol-chloroform RNA extraction

To extract RNA, cells were washed three times with PBS and lysed in 800 μl TRIzol® reagent (phenol; Invitrogen). The lysate was vortexed, 200 μl added, mixed and incubated at room temperature for 2 min. The solution was centrifuged at 13,200 rpm at 4ºC for 15 min and the clear supernatant was removed and mixed with 500 μl isopropanol. Incubation for 10 min at room temperature precipitated the RNA, which was pelleted at 13,200 rpm at 4ºC for 15 min. The pellet was washed with 1 ml 75% ethanol and centrifuged at 7,500 rpm at 4ºC for 5 min. After evaporation of all ethanol traces in a bath at 60ºC, the RNA pellet was resuspended in water and dissolved for 10 min at 60ºC prior to quantification.

7.2. Quantitative RT-PCR

RNA was retrotranscribed using oligo dT and the Transcriptor First Strand cDNA Synthesis Kit (Roche) following manufacturer's instructions. Analyses were carried out in triplicates with 50-100 ng of cDNA using the LightCycler 480 Real Time PCR System (Roche). The primers used for quantitative RT-PCR are indicated in *Table M.2*.

8. Chromatin immunoprecipitation (ChIP)

Cells were crosslinked in 1% formaldehyde for 10 min at 37ºC. Crosslinking was stopped by adding glycine to a final concentration of 0.125 M for 2 min at room temperature. Cell monolayers were scraped in cold Soft-lysis buffer and incubated 20 min on ice. Nuclei pellets were lysed with SDS-lysis buffer and extracts were sonicated to generate 200- to 750-bp DNA fragments. After protein quantification, 100 μg were used for histone H3 ChIPs while 1000 μg was used with transcription factor ChIPs. After taking corresponding volume from the samples, they were diluted 1:10 with Dilution buffer, and immunoprecipitation was done by rotation overnight at 4ºC with primary antibody or irrelevant IgGs, followed by 3 hr incubation with BSA blocked protein A or protein G agarose beads (40 μl/sample; Diagenode). Precipitated samples were then washed three times with Low salt washing buffer and with High salt washing buffer and twice with LiCl buffer using columns. Samples were then treated with 100 μl Elution buffer for 1 hr at 37 ºC and incubated at 65ºC overnight after addition of NaCl at a final concentration of 200 mM to reverse formaldehyde crosslinking. After proteinase K treatment for 1 hr at 55ºC, DNA was purified with MinElute PCR purification kit from Qiagen and eluted in MilliQ water (80 μl – 180 μl). Promoter regions were detected by quantitative PCR and the ChIP results were quantified relative to the input amount.

To detect histone H3 and TAF10 oxidation, ChIP assays were performed as described. After sample sonication, hydrazide-biotin was added at 5 mM final concentration to the supernatants and incubated at 25ºC for 2 hr. After sample dilution 1:10 with Dilution buffer, samples were subjected to immunoprecipitation with anti-histone H3 or anti-TAF10 antibodies overnight at 4ºC followed by 3 hr incubation with protein A or G agarose beads (40 μl/sample; Diagenode). Precipitates were re-extracted with 100 μl Elution buffer for 1 hr at 37ºC, and then diluted again 1:10 dilution with dilution buffer, and re-immunoprecipitated with streptavidin-magnetic beads for 45 min at 4ºC. Samples were then washed three times with Low salt washing buffer and with High salt washing buffer and twice with LiCl. Finally samples were eluted with 100 μl Elution buffer for 1 hr at 37ºC and incubated at 65ºC overnight after addition of NaCl at a final concentration of 200 mM to reverse formaldehyde crosslinking. Results were quantified by taking into account the total amount of histone H3 or TAF10 immunoprecipitated in each condition.

Soft lysis buffer

50 mM Tris-HCl pH 8.0
2 mM EDTA
0.1% Nonidet P-40
10% Glycerol
Protease and phosphatase inhibitors

SDS lysis buffer

50 mM Tris-HCl pH 8.0
10 mM EDTA
1% SDS

Dilution buffer

0.01% SDS

1.1% Triton X-100

1.2 mM EDTA

167 mM NaCl

16.7 mM Tris-HCl pH 8.0

Elution buffer

1% SDS

100 mM Na_2CO_3

High salt buffer

0.1% SDS

1% Triton X-100

2 mM EDTA

20 mM Tris-HCl pH 8.0

500 mM NaCl

Low salt buffer

0.1% SDS

1% Triton X-100

2 mM EDTA

20 mM Tris-HCl pH 8.0

150 mM NaCl

LiCl buffer

250 mM LiCl

1% Nonidet P-40

1% Sodium deoxycholate

1 mM EDTA

10 mM Tris-HCl pH 8.0

9. Alkaline Phosphatase and Clonogenic assays

For alkaline phosphatase assays, ESCs were plated in a 6-well plate for 96 hr, and AP assays were performed with the alkaline phosphatase detection kit (Millipore) following the manufacturer's instructions.

For clonogenic assays, ESCs shControl and shLoxl2 were plated in a 6-well plate at a clonogenic density (1000, 5000 and 10,000 cells/well) for 1 day with puromycin (2 μg/μl). Cells were treated with ATRA 1 mM for 72 hr maintaining puromycin selection. Media was replaced every 24 hr. For the assay, media was removed and cells were washed with PBS. Then, cells were fixed and stained with a mixture of 6% glutaraldehyde and 0.5% crystal violet for at least 30 min. Plates were then washed with tap water, not by putting the plates under the running tap but by immersing the dishes in water carefully. Finally, plates with colonies were left to dry in normal air at room temperature.

10. Mass spectrometry (MS) analysis

10.1. Sample preparation and analysis

Samples were reduced with dithiothreitol (33.3 nM, 1 h, 37°C) and alkylated in the dark with iodoacetamide (66.7 nM, 30 min, 25 ºC). The protein mixture was then diluted 6 times with 200 mM ABC and digested with 6 µg of trypsin (Promega, cat # V5113) overnight at 37 ºC. Samples were acidified with formic acid and cleaned up on a homemade Empore C18 column (3M, St. Paul, MN, USA)[234].

Samples were analysed using a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an Agilent Technologies 1200 Series (CA, USA). Peptides were loaded onto C18 Zorbax precolumn (Agilent Technologies, cat #5065-9913) and were separated by reversed-phase chromatography using a 12-cm column with an inner diameter of 75 μm, packed with 5 μm C18 particles (Nikkyo Technos Co., Ltd. Japan). Chromatographic gradients started at 97% buffer A and 3% buffer B with a flow rate of 300 nl/min, and gradually increased to 90% buffer A and 10% buffer B in 1 min, and to 65% buffer A / 35% buffer B in 60 min. After each analysis, precolumn and column were washed for 10 min with 10% buffer A / 90% buffer B. Buffer A: 0.1% formic acid in water. Buffer B: 0.1% formic acid in Acetonitrile.

The mass spectrometer was operated in positive ionization mode with nanospray voltage set at 2.5 kV and source temperature at 200 °C. Ultramark 1621 for the FT mass analyser was used for external calibration prior the analyses. Moreover, an internal calibration was also performed using the background polysiloxane ion signal at m/z 445.1200. The instrument was operated in DDA mode and full MS scans with 1 micro scans at resolution of 60.000 were used over a mass range of m/z 350-2000 with detection in the Orbitrap. Auto gain control (AGC) was set to 1E6, dynamic exclusion (60 seconds) and charge state filtering disqualifying singly charged peptides was activated. In each cycle of DDA analysis, following each survey scan the top ten most intense ions with multiple charged ions above a threshold ion count of 5000 were selected for fragmentation at normalized collision energy of 35%. Fragment ion spectra produced via collision-induced dissociation (CID) were acquired in the Ion Trap, AGC was set to 5e4, isolation window of 2.0 m/z, activation time of 0.1ms and maximum injection time of 100 ms was used. All data were acquired with Xcalibur software v2.2.

10.2. Data analysis

Proteome Discoverer software suite (v1.4.1.14, Thermo Fisher Scientific) and the Mascot search engine (v2.3, Matrix

Science[235]) were used for peptide identification. The data were searched against a SwissProt database containing entries corresponding to Homo sapiens (version of January 2012), a list of common contaminants and all the corresponding decoy entries. Trypsin was chosen as enzyme and a maximum of three miscleavages were allowed. Carbamidomethylation (C) was set as a fixed modification, whereas oxidation (M) and (K) and acetylation (N-terminal), phosphorylation (S), dimethyl (K), trimethyl (K) and methyl (K) were used as variable modifications. Searches were performed using a peptide tolerance of 7 ppm, a product ion tolerance of 0.5 Da. Peptides have been filtered based on Xcorr (z=2 xcorr=0.9, z=3 xcorr=1.2, z=4,5,6 xcorr=1.5). Resulting data files were filtered for false discovery rate (FDR) < 5 %.

11. Gene expression microarray analysis

We measured gene expression levels of shControl-ESCs and shLoxl2-ESCs treated or not with RA. For microarrays analysis, amplification, labelling, and hybridizations were performed according to protocol GeneChip WT PLUS Reagent kit and hybridized to the GeneChip Mouse Gene 2.0 ST Array (Affymetrix) in a GeneChip Hybridization Oven 640. Washing and scanning were performed using the Expression Wash, Stain, and Scan Kit and the GeneChip System of Affymetrix (GeneChip Fluidics Station 450 and GeneChip Scanner 3000 7G). After quality control of raw data, data were background corrected, quantile-normalized, and summarized to a gene level using the robust multichip average (RMA)[236], resulting in a total of 34390 transcript clusters, excluding controls, which roughly correspond to genes or other mRNAs as miRNAs or lincRNAs. Linear Models for Microarray (LIMMA)[237], a moderated t-statistics model, was used to detect differentially expressed genes between the conditions. Correction for multiple comparisons was performed using false discovery rate. Genes with $p < 0.05$ and with an absolute fold change (FC) value above 1.5 were selected as significant. Ingenuity Pathway Analysis (Ingenuity Systems) was used to functionally analyse the results. Comparison with published results was performed as follows. Data were downloaded from GEO under the accession number GSE33644. This experiment was

performed in the two-color Agilent based microarray A-UMCUM44K-1.0, following a dye swap design against a common reference. Arrays were background corrected and normalized using the loess method. Analysis was assessed using the same methodology as previously described and results of both studies were compared in terms of gene coincidences. All analyses were performed in R (v 3.1.1) with packages aroma.affymetrix, limma, Biobase, Vennerable, and XLConnect.

12. Zebrafish

12.1. Fish maintenance and stains

AB wild-type zebrafish were maintained at the PRBB Animal Facility under standard conditions[238]. Embryos were developed in system water containing methylene blue in an incubator at 28.5ºC. Embryonic stages are given as hours post-fertilization (hpf) at 28.5ºC.

12.2. Whole-mount *in situ* hybridization

Zebrafish full gene names are lowercase italic (ZFIN Nomenclature Guidelines). Probes for detecting loxl2a and loxl2b gene expression were synthesized by PCR. T7 RNA polymerase promoter was introduced at the 50 end of each gene-specific oligonucleotide reverse primer, enabling direct in vitro transcription of the purified PCR products. The primer sequences are the following: loxl2a forward primer, 5'-TCAGTTGGTGGAGCAGACC-3'; loxl2a reverse primer, 5'-TTGTAATACGACTCACTATAGGCAGCGACGTCATAGTTTGGA-3'; loxl2b forward primer, 5'-CCTAAAGGTGGAGGACGATTC-3'; and loxl2b reverse primer, 5'-TTGTAATACGACTCACTATAGGACCCACTGGCAATCAAT

GTC-3’. Additional probes used were previously described for Sox2[239] and NeuroD[240].

Whole-mount in situ hybridization was performed as previously described[241]. Embryos were isolated at the desired developmental stages essentially as described[238]. Dechorionated zebrafish embryos were fixed in 4% paraformaldehyde (PFA) overnight at 4ºC, dehydrated in a methanol series, rehydrated again and permeabilized with 10 mg/mL proteinase K (Sigma) at room temperature for 5–10 min, depending on their stage. Digoxigenin-labeled probes were hybridized overnight at 70ºC, detected using anti-digoxigenin-AP antibody at 1:2000 dilution (Roche), and developed with NBT/BCIP (Roche). Embryos were postfixed overnight in 4% PFA and used for imaging mounted in 100% glycerol.

12.3. Morpholinos

Morpholino oligomers (MOs) targeting splice sites of loxl2 exons and a standard control MO were obtained from Gene Tools. Embryos were injected into the yolk at the one-cell stage with 2.5 ng of loxl2b and loxl2a morpholino mix, together with GFP mRNA to assess the efficiency of injection. MO for loxl2b was previously described with sequence: 5’-GATCTGGAGCAGCTAGAAAAAACAA-3’ (targeted to intron9

and exon10 junction)[185]. The sequences of the loxl2a-MO and the standard control MO were 5'-TCAATAAGCTTTTGGCCTACCTTCT-3' (targeted to exon1 and intron1 junction) and 5'-CCTCTTACCTCAGTTACAATTTATA-3', respectively. Abnormal splicing caused by the loxl2a-MO was detected by RT-PCR of RNA from injected embryos using primers to exons flanking the putative splice site. The sequences of the primer F1 and R2 were 5'-CTCAGGCCCAATCTGAACTC-3' and 5'-TGATGAACCTGTGACCAGGA-3', respectively.

12.4. Image acquisition

Pictures were acquired in a Leica DRM microscope using a Leica DFC300 FX camera and the Leica IM50 software. Adobe Photoshop 7.0.1 software was used for photograph editing.

13. Antibodies used

Protein	Species	Provider	Reference	Dilution
FLAG	rabbit pAb	Sigma	F7425	WB 1:2000
HA	rabbit pAb	Sigma	H6908	IP 1:250
	rat pAb	Roche	65850900	WB 1:2000
Pyruvate Kinase	mouse mAb	Chemicon	AB1235	WB 1:4000
LOXL2	mouse mAb	Abcam	ab55470	ChIP 1:50
	rabbit pAb	Novus	NBP1-32954	WB 1:1000
TAF10	rabbit pAb	NeoBiotech	NB-01-0132	WB 1:2000 IP 1:100
	mouse mAb	Abnova	H00006881-D01	WB 1:2000 IP 1:100 ChIP 1:200
TBP	rabbit pAb	Abcam	ab63766	WB 1:2000 IP 1:100
	mouse mAb	Abcam	ab51841	WB 1:2000 IP 1:100 ChIP 1:200
TAF1	rabbit pAb	Abcam	ab28450	ChIP 1:200
TAF11		Abcam		
	mouse mAb	NeoBiotech	AT4138a	WB 1:5000
TAF13		NeoBiotech	NB-01-0137	WB 1:1000
RNAPII				ChIP 1:200
H3	rabbit pAb	Abcam	ab1791	ChIP 1:300

Table MM.1. Antibodies and their applications. The antibodies used in this study, their commercial information and dilution for use are described in detail.

14. Primers used

	mRNA-qPCR		ChIP-qPCR
PLK1	TCCAAGCCCTCGGAGCGT	PLK1	GGTAGGGCCTGAACGTTTGC
	ACAGAGCTGATACCCAAGGCCGT		CGTGGATGCGGAGACCC
ERF1	GGTTGGCTGGGGGTGCAGT	ERF1	GCTTCTTCGCGGCTTCTCAA
	GTTCCTCCAGCTCTGACGTGCC		GGTCGTCGTCTTTTCAGTCCATTC
HOXA1	CCCAAAACAGGGAAAGTTGG	HOXA1	CTCTTCTTCGCTCCAGCACTCC
	TTCTCACGTTTCTTTTGCTTC		TGGCATTTAAATCCCCGGC
CCNE1	CCATCATGCCGAGGGAG	CCNE1	AGGTCTTCAGAGAGCCAGGA
	ATTGTCCCAAGGCTGGCTC		GGCCTAGAACCAAGGCTTC
HPRT	GGCCAGACTTTGTTGGATTTG	HPRT	ATTCACGCGATGACTGGA
	TGCGCTCATCTTAGGCTTTGT		AGGCTCACTAGGTAGCCGTG
Pumilio	CGGTCGTCCTGAGGATAAAA		
	CGTACGTGAGGCGTGAGTAA		

Table MM.2. Human primers used for mRNA and ChIP analysis. The sequences of the different human primers used in this study are listed. The first primer is forward and the second reverse. All of them are shown in 5' to 3' direction.

mRNA-qPCR		ChIP-qPCR	
Nanog	AGGCTGATTTGGTTGGTGTC	Nanog	TAAGCTTTCCCTCCCTCCC
	CCAGGAAGACCCACACTCAT		TACCCTACCCACCCCCTATTC
Sox2	CTGCAGTACAACTCCATGACCAG	Sox2	GCTCTGCCAGCTTCCTGAAAT
	GGACTTGACCACAGAGCCCAT		ACGCCTGTTCGAAGGAAGTG
Klf4	GTTTTTAATCTTCGTTGACTT	Klf4	CTGAGTCCAAGAGCGTGCAG
	CAGCCATGTCAGACTCGC		GATTCAATATAAACCGGCGATGTC
Oct4	GAGGAGTCCCAGGACATGAA	Oct4	AACTGAGGAGTGGCCCCAG
	AGATGGTGGTCTGGCTGAAC		CCAGGAGGCCTTCATTTTCAA
Loxl2	GCGAAGGCCCCATCTGGTTG	Rps28	CGCAGCTTTTGGATAACGC
	GTGCTTGCAGTCAGTGACACCCC		ATCTGGCGGAGAGGAGTCAC
Taf10	CACACCTACGATCCCAGATGC	intraNanog	ACTCTGCAGGTTAAGACCTGG
	AGGCTGTGCCCTTCATTTTG		ATCCATACCCAACAGCCATAG
Nestin	AAAGTTCCAGCTGGCTGTGG	interNanog	ACAGGCATCGTAGCAGCA
	AGAGTTCTCAGCCTCCAGCA		GCCCTTCTACCCACTGAGG
NeuroG1	CAGTAGTCCCTCGGCTTCAG	intraSox2	TCTGTGGTCAAGTCCGAGG
	GAAAGGAGAAAAGGGGATCG		GGGCAGTGTGCCGTTAAT
HoxA1	TTCTCCAGCGCAGACCTTTG	interSox2	GGGAGACGCTCAGGTTCA
	GCACTGCGTTGGGTTGACC		TGGGTAGCGCTGCAGTTTAT
HoxA5	TACCCCTGGATGCGCAAG	intraKlf4	TGGCAAGCGCTACAATCAT
	CAGCTCCAGGGTCTGGTAGC		ATCCATCCAGTATCAGACCCC
HoxA7	GAAGCCAGTTTCCGCATCTACC	interKlf4	GCCAACTGATTGGAGGGC
	CTTCTCCAGTTCCAGCGTCTG		AAGACAAAGCTTCAGCCGGT
Txnrd3	TAACCCTTCGGGAGAAAGG	intraOct4	TGCTCCCTTAGCACAATCC
	AGAGAACAGGTCGTCACTCG		CTCCACCTCACACGGTTCT
Myh9	TGTGGAAGGATGTGGATCG	interOct4	TTCCCTTGGGGGTATGG
	CGGCCTTCTTCTCATGGTT		CTGGGGTTGATTCCCAAGT

	mRNA-qPCR	ChIP-qPCR
Dppa3	TGACTGCTAATTGGGTCTTGG	
	TTTCTGGTTGTAGGACGTCTGT	
Trim8	TGGACAGGACCCAGACCT	
	ACCACAAGGGTACAGGGGTA	
Apoe	ACATTGCTGACAGGATGCC	
	TCCATCAGTGCCGTCAGTT	
Cdyl2	ATCATGAAGGAAGTCCGGC	
	GCCTTCACGAAGTCCCTG	
Aoah	ATCCCAGAAGTGGTCGTAACA	
	TCCCAGTTGTCTGGCCTT	
Rpo	TTCATTGTGGGAGCAGAC	
	CAGCAGTTTCTCCAGAGC	

Table MM.3 Mouse primers used for mRNA and ChIP analysis. The sequences of the different mouse primers used in this study are listed. The first primer is forward and the second reverse. All of them are shown in 5' to 3' direction.

ANNEX

1. Putative LOXL2 substrates

Accession	Description	ΣCoverage	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs
Q06136	KDSR	2,11	1	1	1
P27635	RPL10	19,16	2	2	2
Q5T0R7	CAP1	7,47	1	1	1
Q4LEZ3	AARD	7,10	1	1	1
P02763	ORM1	7,46	1	1	2
P02765	AHSG	12,26	2	2	2
P01023	A2M	15,26	15	15	16
Q5T3N1	ANXA1	22,06	3	3	3
P03973	SLPI	20,45	2	2	2
P02647	APOA1	28,09	7	7	9
P02652	APOA2	10,00	1	1	1
P17213	BPI	3,08	1	1	1
P61769	B2M	18,49	1	1	1
P04040	CAT	2,47	1	1	1
P00450	CP	3,29	2	2	3
P01024	C3	8,96	10	10	10
P31146	CORO1A	2,60	1	1	1
P81605	DCD	12,73	1	1	1
Q5T0H9	GSN	8,87	2	2	2
P06744	GPI	5,73	2	2	2
P00738	HP	23,65	7	7	7
P69905	HBA1	21,83	2	2	3
P02790	HPX	6,71	3	3	3
P04196	HRG	2,67	1	1	1
Q9XRX5	HHLA3	7,89	1	1	1
P16403	Histone H1.2	25,82	2	5	17
P16402	Histone H1.3	19,91	1	5	11
P10412	Histone H1.4	25,57	2	6	18
P68431	Histone H3.1	23,53	1	1	1
P62847-2	RPS24	11,54	1	1	1
P62913-2	RPL11	7,91	1	1	1
P01009-2	SERPINA1	6,69	2	2	2
P49959-2	MRE11A	3,38	1	1	1
P02671-2	FGA	3,88	2	2	2
Q5VWN6-2	FAM208B	0,86	1	1	1
Q86UC2-2	RSPH3	2,59	1	1	1
Q12923-2	PTPN13	1,57	2	2	2
Q32MH5-2	KIAA1370	0,83	1	1	1

Accession	Description	ΣCoverage	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs
P02774-2	VTDB	6,25	2	2	2
O60716-32	CTNND1	3,05	1	1	1
P11171-6	EPB41	2,83	1	1	1
Q9UGM3-7	DMBT1	11,33	2	2	2
P15924-2	DSP	0,84	1	1	1
P02679-2	FGG	9,38	3	3	4
O00161-2	SNAP23	8,86	1	1	1
P61626	LYZ	45,95	5	5	6
Q9Y4K0	LOXL2	2,58	1	1	1
P14780	MMP9	5,94	3	3	3
P41218	MNDA	4,18	1	1	1
Q9NQX4	MYO5C	0,52	1	1	1
P48681	NES	2,59	2	2	3
P59665	DEFA1	9,57	1	1	1
P80188	LCN2	16,16	2	2	2
O75594	PGLYRP1	11,73	1	1	1
P32119	PRDX2	5,56	1	1	1
Q7RTV0	PHF5A	11,82	1	1	1
P42356	PI4KA	0,98	1	1	1
Q9NWQ8	PAG1	6,94	1	1	1
P00747	PLG	2,59	2	2	2
P01833	PIGR	25,39	13	13	16
Q7L014	DDX46	1,45	1	1	2
P07737	PFN1	21,43	2	2	2
P12273	PIP	4,11	1	1	1
P31949	S100A11	15,24	1	1	1
P80511	S100A12	21,74	1	1	1
P06703	S100A6	18,89	1	1	1
P05109	S100A8	23,66	2	2	2
P06702	S100A9	41,23	4	4	4
P25815	S100P	34,74	2	2	2
Q9UNW9	NOVA2	3,66	1	1	1
P02787	TF	21,35	12	12	16
P02768	ALB	67,00	37	41	65
P62328	TMSB4X	31,82	1	1	1
F2Z393	TALDO1	3,46	1	1	1

Accession	Description	ΣCoverage	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs
B4DS69	ARFIP1	5,18	1	1	1
E7EUG0	ARGLU1	3,59	1	1	1
F5H6Q0	ARHGDIB	22,32	1	1	1
F5H0J0	ATP1A3	2,64	1	1	1
E7EQ46	BPIL1	6,24	1	1	1
A8MWK3	CDH2	2,06	1	1	3
F5H3E5	DDX55	1,98	1	1	1
F5H521	DHX8	1,99	1	1	1
A6NI74	ENO4	3,31	1	1	1
D6REL8	FGB	37,87	7	7	8
F5H407	FLII	3,48	2	2	2
A8MX94	GSTP1	24,14	2	2	2
F8W6P5	HBB	14,44	1	1	1
C9JQA7	HMG20B	4,93	1	1	1
D6RD17	IGJ	16,46	3	3	3
F5H7Y8	ITIH1	5,14	2	2	3
F5GXQ1	ITIH4	2,86	1	1	1
C9JEX1	KNG1	3,32	1	1	1
B4DUA0	LCP1	14,80	2	2	3
F5GXY2	LDHA	10,83	1	1	3
B7Z4X2	LTF	3,30	2	2	2
E9PCI8	MSH2	8,33	1	1	1
C9J9C7	PARP3	5,12	1	1	6
F5GXG4	PGAM5	7,14	1	1	1
C9JK77	PLA2G4E	1,19	1	1	1
F5H8B6	PRTN3	9,77	2	2	3
F5H4N4	PXMP2	6,25	1	1	1
E7ESD9	RAD50	2,90	1	1	1
C9JD32	RPL23	21,98	1	1	1
A8MUD9	RPL7	9,13	1	1	1
B4E3L0	SMU1	4,83	1	1	1
E7EUL7	SSFA2	1,69	1	1	2
B3KSI4	TKT	13,15	4	4	4
F8WB30	TOM1	64,00	1	1	1

Accession	Description	ΣCoverage	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs
Q12962	TAF10	27,52	3	3	10
Q15544	TAF11	7,58	1	1	1
Q16514-2	TAF12	11,45	2	2	2
Q15543	TAF13	8,87	1	1	1
Q15545	TAF7	4,58	1	1	1
E9PMR3	WRAP53	6,27	1	1	1
E5RGE1	YWHAZ	26,92	1	1	2
Q9UPR6	ZFR2	1,60	1	1	1
Q96DA0	ZG16B	16,83	3	3	3
F5GXY2	LDHA	10,83	1	1	3
B7Z4X2	LTF	3,30	2	2	2
E9PCI8	MSH2	8,33	1	1	1
C9J9C7	PARP3	5,12	1	1	6
F5GXG4	PGAM5	7,14	1	1	1
C9JK77	PLA2G4E	1,19	1	1	1
F5H8B6	PRTN3	9,77	2	2	3
F5H4N4	PXMP2	6,25	1	1	1
E7ESD9	RAD50	2,90	1	1	1
C9JD32	RPL23	21,98	1	1	1
A8MUD9	RPL7	9,13	1	1	1
B4E3L0	SMU1	4,83	1	1	1
E7EUL7	SSFA2	1,69	1	1	2
B3KSI4	TKT	13,15	4	4	4
F8WB30	TOM1	64,00	1	1	1
E9PMR3	WRAP53	6,27	1	1	1
E5RGE1	YWHAZ	26,92	1	1	2
Q9UPR6	ZFR2	1,60	1	1	1
Q96DA0	ZG16B	16,83	3	3	3

Table A.1. Putative LOXL2 substrates. List of proteins identified by MS as LOXL2 substrate candidates. Coverage of the sequence, number of aminoacids, number of unique peptides, total peptides and number of peptide spectrum matches (PSM).

2. KEGG pathways of LOXL2 putative substrates

GO_ID	Term	Genes	Genes nº	p-value
4610	Complement and coagulation cascades	A2M FGB KNG1 FGA FGG SERPINA1 PLG C3	8	3.05E-9
3022	Basal transcription factors	TAF7 TAF11 TAF12 TAF13 TAF10	5	3.77E-6
30	Pentose phosphate pathway	TALDO1 GPI TKT	3	3.59E-4
4611	Platelet activation	PLA2G4E FGB FGG SNAP23 FGA	5	5.98E-4
3010	Ribosome	RPL7 RPL10 RPL23 RPS24 RPL11	5	6.42E-4
3450	Non-homologous end-joining	MRE11A RAD50	2	2.15E-3
5150	Staphylococcus aureus infection	FGG PLG C3	3	2.61E-3
3440	Homologous recombination	RAD50 MRE11A	2	9.21E-3
1200	Carbon metabolism	GPI TKT TALDO1	3	1.77E-2
3320	PPAR signalling pathway	APOA1 APOA2	2	5.19E-2
1230	Biosynthesis of amino acids	TKT TALDO1	2	5.46E-2
4146	Peroxisome	CAT PXMP2	2	6.61E-2

Table A.2. KEGG pathway analysis of LOXL2 putative substrates. Candidates were searched at KEGG pathway database to check in which processes were involved. Significantly enriched pathways are shown.

3. LOXL2 substrate candidates with SET7/9 recognition motif

Accession	Description	Site	Sequence	Surface accessibility
P01024	C3	K427	LSITVRT**K**KQELSEA	7,7703
		K1353	TMYHAKA**K**DQLTCNK	2,2084
Q9XRX5	HHLA3	K78	CRRGMKA**K**QLKAEAG	2,6446
P16403	Histone H1.2	K159	KKTPKKA**K**KPAAATV	2,8926
P16402	Histone H1.3	K25	TPVKKKA**K**KAGATAG	1,8512
		K122	GEGKPKA**K**KAGAAKP	1,8512
		K214	KPKVTKA**K**KAAPKKK	1,8898
P10412	Histone H1.4	K121	GEAKPKA**K**KAGAAKA	1,8512
		K129	KAGAAKA**K**KPAGAAK	2,8926
		K159	KKTPKKA**K**KPAAAAG	2,8926
Q5VWN6-2	FAM208B	K1397	KSFFVRT**K**NLLRKGG	1,4713
Q12923-2	PTPN13	K1114	DSNQSKT**K**KPGISDV	4,0479
		K1449	NEMSDKS**K**KQCKSPS	2,2803
P11171-6	EPB41	K160	GVDLHKA**K**DLEGVDI	2,2084
P15924-2	DSP	K443	QNLVNKS**K**KIVQLKP	1,278
P41218	MNDA	K397	FIKVIKA**K**KNKEGPM	6,1842
P00747	PLG	K119	RGTMSKT**K**NGITCQK	1,5324
Q7L014	DDX46	K112	SSPGNKS**K**KTENRSR	6,1393
Q9UNW9	NOVA2	K72	TIKLSKS**K**DFYPGTT	2,783
		K131	TMNPDRA**K**QAKLIVP	3,1729
Q12962	TAF10	K189	GSSRSKS**K**DRKYTLT	8,0343
F5H521	DHX8	K172	AEHRDRT**K**KKKRSRS	10,3615
C9J9C7	PARP3	K134	KKFREKT**K**NNWAERD	3,8788
E7ESD9	RAD50	K107	MVCTQKS**K**KTEFKTL	6,1393
E7EUL7	SSFA2	K227	GRSLLKS**K**DLLKQRY	1,395
B3KSI4	TKT	K416	AKVVLKS**K**DDQVTVI	5,9322

Table A.3. LOXL2 substrate candidates with SET7/9 recognition motif. LOXL2 putative substrates presenting a SET7/9 recognition motif. Methylation candidate lysines, surrounding sequence and surface accessibility are shown.

4. Posttranslational modifications identified in TAF10

Sequence	Modifications	PSMs	A Area	B Area	C Area	A XCorr	B XCorr	C XCorr
DRKYTLTMEDLTPALSEYGINVK	K3(Oxidation) M8(Oxidation) S16(Phospho) K23(Methyl)	8	0,00E+00	5,41E+07	5,15E+07		3,19	2,5
FISDIANDALQHCKMKGTASGSSRSK	C13(Carbamidomethyl) K14(Oxidation) M15(Oxidation) K16(Oxidation) K26(Trimethyl)	1	0,00E+00	7,26E+06	0,00E+00		2	
GTASGSSRSK	K10(Dimethyl)	2	0,00E+00	5,59E+06	1,43E+07		1,76	1,66
GTASGSSRSKSK	N-Term(Acetyl) S4(Phospho) K10(Trimethyl) K12(Oxidation)	2	0,00E+00	6,83E+06	9,02E+06		2,58	2,84
GTASGSSRSKSK	S4(Phospho) S9(Phospho) K10(Dimethyl) S11(Phospho)	1	0,00E+00	0,00E+00	2,01E+09			1,42
GTASGSSRSKSKDRK	N-Term(Acetyl) S11(Phospho) K15(Oxidation)	2	0,00E+00	5,21E+06	9,42E+06		2,47	2,22
MKGTASGSSR	N-Term(Acetyl) K2(Oxidation)	1	0,00E+00	1,48E+07	0,00E+00		2,04	
MKGTASGSSR	K2(Methyl) S6(Phospho)	1	0,00E+00	0,00E+00	7,73E+06			1,33
MKGTASGSSRSK	N-Term(Acetyl) K2(Trimethyl) S6(Phospho)	1	9,02E+05	0,00E+00	0,00E+00	1,85		
MKGTASGSSRSK	M1(Oxidation) S11(Phospho) K12(Trimethyl)	5	0,00E+00	1,40E+07	4,93E+07		2,27	2,71

Sequence	Modifications	PSMs	A Area	B Area	C Area	A XCorr	B XCorr	C XCorr
MKGTASGSSRSK	N-Term(Acetyl) K2(Trimethyl) S11(Phospho) K12(Oxidation)	1	0,00E+00	9,96E+06	0,00E+00		1,88	
MKGTASGSSRSK	M1(Oxidation) K2(Methyl) S6(Phospho)	1	0,00E+00	0,00E+00	1,28E+07			1,82
MKGTASGSSRSKSK	K2(Methyl) S6(Phospho) S9(Phospho) K12(Dimethyl)	1	5,14E+05	0,00E+00	0,00E+00	2,84		
MKGTASGSSRSKSK	M1(Oxidation) K2(Oxidation) K12(Methyl) S13(Phospho)	1	0,00E+00	0,00E+00	2,56E+07			1,45
MKGTASGSSRSKSK	M1(Oxidation) K2(Oxidation) S11(Phospho) K12(Trimethyl)	1	0,00E+00	0,00E+00	1,21E+07			2,44
MKGTASGSSRSKSKDR	N-Term(Acetyl) S6(Phospho) K12(Oxidation) K14(Trimethyl)	2	2,32E+07	0,00E+00	9,56E+07	2,06		2,33

Table A.4. Posttranslational modifications identified in TAF10. Posttranslational modifications identified in TAF10 by mass spectrometry are shown. Type of modification, target residue, surrounding sequence, peptide spectrum matches, peptide area, and cross correlation (xcorr) are depicted.

5. Gene expression profile comparison

TAF5 dependent genes downregulated +RA	TAF5 dependent genes downregulated in a LOXL2 dependent manner +RA
Symbol	Symbol
Ddx58	Aoah
Slc27a2	Trim8
8430410A17Rik	Gm129
Enox1	Apoe
Pcolce2	Nid1
Slc7a3	4930519F16Rik
Foxd3	Eed
Ntn1	Cdyl2
1190003J15Rik	Trim25
Elmo1	Trib3
Ildr1	Triml1
Gtsf1l	Rabggtb
Nr0b1	Aff1
Gcnt2	Hexa
Kirrel2	Klf8
Slc17a9	Sod2
Mcam	Zfp740
Exoc3l	Nusap1
Trim25	Chac1
Gpa33	Fry
1700061G19Rik	
Aoah	
Chrna9	
Atp1b1	
Zbtb8a	
Sod2	
Mybl2	
Prr13	
Spry4	
Sall1	
Fry	
Morc1	
Cbr3	
Syk	
Hck	
Pdk1	
Zfp740	

TAF5 dependent genes downregulated +RA	TAF5 dependent genes downregulated in a LOXL2 dependent manner +RA
Symbol	Symbol
2610318N02Rik	
Nphs1	
Csrp2	
Nid1	
Gm129	
Ndrg1	
Dnmt3l	
Zfp459	
Dab1	
Trim8	
Pla1a	
Aff1	
Pdzd2	
Socs2	
Mcf2	
Lrrc34	
Ifitm3	
sep-01	
AU018091	
Vangl1	
Apoe	
Pcyt1b	
Cdyl2	
BC028528	
Tns3	
Aass	
Hpdl	
Bmp4	
Cyp2s1	
Hexa	
Fgf4	
Emp1	
Myof	
Pdgfc	
Klf8	

TAF5 dependent genes downregulated +RA	TAF5 dependent genes downregulated in a LOXL2 dependent manner +RA
Symbol	Symbol
Cltb	
Pcolce	
Pim3	
Spats1	
Aplp1	
Nusap1	
Adam23	
Mical1	
Gli2	
Epas1	
Zfp296	
Eed	
Triml1	
Setx	
Esrrb	
Cth	
Gjb5	
Ly75	
Rabggtb	
Pfkp	
Rmnd5b	
Trib3	
4930519F16Rik	
Gjb3	
Chac1	

Table A.5. Gene expression profile comparison. Left column, comparison of microarray data from TAF5 dependent genes[80] and downregulated genes in control cells upon RA treatment. Right column, comparison of left column subset genes with downregulated genes upon RA treatment in a LOXL2 dependent manner.

REFERENCES

1 Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871 (1974).

2 Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews. Molecular cell biology* **16**, 178-189, doi:10.1038/nrm3941 (2015).

3 Lavelle, C. Pack, unpack, bend, twist, pull, push: the physical side of gene expression. *Current opinion in genetics & development* **25**, 74-84, doi:10.1016/j.gde.2014.01.001 (2014).

4 Palstra, R. J. *et al.* The beta-globin nuclear compartment in development and erythroid differentiation. *Nature genetics* **35**, 190-194, doi:10.1038/ng1244 (2003).

5 Bantignies, F. *et al.* Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. *Cell* **144**, 214-226, doi:10.1016/j.cell.2010.12.026 (2011).

6 Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harbor perspectives in biology* **2**, a003889, doi:10.1101/cshperspect.a003889 (2010).

7 Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular cell* **38**, 603-613, doi:10.1016/j.molcel.2010.03.016 (2010).

8 Chaumeil, J., Le Baccon, P., Wutz, A. & Heard, E. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes & development* **20**, 2223-2237, doi:10.1101/gad.380906 (2006).

9 Therizols, P. *et al.* Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* **346**, 1238-1242, doi:10.1126/science.1259587 (2014).

10 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306-1311, doi:10.1126/science.1067799 (2002).

11 van Berkum, N. L. *et al.* Hi-C: a method to study the three-dimensional architecture of genomes. *Journal of visualized experiments : JoVE*, doi:10.3791/1869 (2010).

12 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

13 Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385, doi:10.1038/nature11049 (2012).

14 Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458-472, doi:10.1016/j.cell.2012.01.010 (2012).

15 Sexton, T. & Cavalli, G. The role of chromosome domains in shaping the functional genome. *Cell* **160**, 1049-1059, doi:10.1016/j.cell.2015.02.040 (2015).

16 Giorgetti, L. *et al.* Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950-963, doi:10.1016/j.cell.2014.03.025 (2014).

17 Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59-64, doi:10.1038/nature12593 (2013).

18 Eissenberg, J. C. & Reuter, G. Cellular mechanism for targeting heterochromatin formation in Drosophila. *International review of cell and molecular biology* **273**, 1-47, doi:10.1016/S1937-6448(08)01801-7 (2009).

19 Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **115**, 110-122, doi:10.1007/s00412-006-0052-x (2006).

20 Flamm, W. G., Walker, P. M. & McCallum, M. Some properties of the single strands isolated from the DNA of the nuclear satellite of the mouse (Mus musculus). *Journal of molecular biology* **40**, 423-443 (1969).

21 Eymery, A. *et al.* A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. *Nucleic acids research* **37**, 6340-6354, doi:10.1093/nar/gkp639 (2009).

22 Henikoff, S. Heterochromatin function in complex genomes. *Biochimica et biophysica acta* **1470**, O1-8 (2000).

23 Allshire, R. C., Nimmo, E. R., Ekwall, K., Javerzat, J. P. & Cranston, G. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes & development* **9**, 218-233 (1995).

24 Kellum, R. & Alberts, B. M. Heterochromatin protein 1 is required for correct chromosome segregation in Drosophila embryos. *Journal of cell science* **108 (Pt 4)**, 1419-1431 (1995).

25 Jia, S., Yamada, T. & Grewal, S. I. Heterochromatin regulates cell type-specific long-range chromatin interactions essential for directed recombination. *Cell* **119**, 469-480, doi:10.1016/j.cell.2004.10.020 (2004).

26 Grewal, S. I. & Elgin, S. C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399-406, doi:10.1038/nature05914 (2007).

27 Maison, C. *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nature genetics* **30**, 329-334, doi:10.1038/ng843 (2002).

28 Valgardsdottir, R. *et al.* Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Molecular biology of the cell* **16**, 2597-2604, doi:10.1091/mbc.E04-12-1078 (2005).

29 Lu, B. Y., Emtage, P. C., Duyf, B. J., Hilliker, A. J. & Eissenberg, J. C. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. *Genetics* **155**, 699-708 (2000).

30 Yasuhara, J. C. & Wakimoto, B. T. Oxymoron no more: the expanding world of heterochromatic genes. *Trends in genetics : TIG* **22**, 330-338, doi:10.1016/j.tig.2006.04.008 (2006).

31 Bonasio, R., Tu, S. & Reinberg, D. Molecular signals of epigenetic states. *Science* **330**, 612-616, doi:10.1126/science.1191078 (2010).

32 Brownell, J. E. & Allis, C. D. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Current opinion in genetics & development* **6**, 176-184 (1996).

33 Musselman, C. A., Lalonde, M. E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature structural & molecular biology* **19**, 1218-1227, doi:10.1038/nsmb.2436 (2012).

34 Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86-90, doi:10.1038/nature04815 (2006).

35 Pena, P. V. *et al.* Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100-103, doi:10.1038/nature04814 (2006).

36 Shi, X. *et al.* ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96-99, doi:10.1038/nature04835 (2006).

37 Fillingham, J., Keogh, M. C. & Krogan, N. J. GammaH2AX and its role in DNA double-strand break repair. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **84**, 568-577, doi:10.1139/o06-072 (2006).

38 Sanders, S. L. *et al.* Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* **119**, 603-614, doi:10.1016/j.cell.2004.11.009 (2004).

39 Doyon, Y. *et al.* ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Molecular cell* **21**, 51-64, doi:10.1016/j.molcel.2005.12.007 (2006).

40 Iizuka, M., Matsui, T., Takisawa, H. & Smith, M. M. Regulation of replication licensing by acetyltransferase Hbo1. *Molecular and cellular biology* **26**, 1098-1108, doi:10.1128/MCB.26.3.1098-1108.2006 (2006).

41 Fischle, W. *et al.* Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-1122, doi:10.1038/nature04219 (2005).

42 Liu, C. L. *et al.* Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS biology* **3**, e328, doi:10.1371/journal.pbio.0030328 (2005).

43 Day, J. J. & Sweatt, J. D. Epigenetic treatments for cognitive impairments. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **37**, 247-260, doi:10.1038/npp.2011.85 (2012).

44 Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nature structural & molecular biology* **20**, 1147-1155, doi:10.1038/nsmb.2669 (2013).

45 Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840, doi:10.1126/science.1175371 (2009).

46 Chuikov, S. *et al.* Regulation of p53 activity through lysine methylation. *Nature* **432**, 353-360, doi:10.1038/nature03117 (2004).

47 Huang, J. *et al.* G9a and Glp methylate lysine 373 in the tumor suppressor p53. *The Journal of biological chemistry* **285**, 9636-9641, doi:10.1074/jbc.M109.062588 (2010).

48 Huang, J. *et al.* Repression of p53 activity by Smyd2-mediated methylation. *Nature* **444**, 629-632, doi:10.1038/nature05287 (2006).

49 Lu, T. *et al.* Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 46-51, doi:10.1073/pnas.0912493107 (2010).

50 Yang, J. *et al.* Reversible methylation of promoter-bound STAT3 by histone-modifying enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21499-21504, doi:10.1073/pnas.1016147107 (2010).

51 Cho, H. S. *et al.* Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. *Cancer research* **71**, 655-660, doi:10.1158/0008-5472.CAN-10-2446 (2011).

52 Wilson, G. G. & Murray, N. E. Restriction and modification systems. *Annual review of genetics* **25**, 585-627, doi:10.1146/annurev.ge.25.120191.003101 (1991).

53 Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annual review of biochemistry* **74**, 481-514, doi:10.1146/annurev.biochem.74.010904.153721 (2005).

54 Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322, doi:10.1038/nature08514 (2009).

55 Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11995-11999 (1993).

56 Chuang, L. S. *et al.* Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277**, 1996-2000 (1997).

57 Jaenisch, R., Harbers, K., Jahner, D., Stewart, C. & Stuhlmann, H. DNA methylation, retroviruses, and embryogenesis. *Journal of cellular biochemistry* **20**, 331-336, doi:10.1002/jcb.240200403 (1982).

58 Issa, J. P. CpG-island methylation in aging and cancer. *Current topics in microbiology and immunology* **249**, 101-118 (2000).

59 Kondo, T. *et al.* Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Human molecular genetics* **9**, 597-604 (2000).

60 Verona, R. I., Mann, M. R. & Bartolomei, M. S. Genomic imprinting: intricacies of epigenetic regulation in clusters. *Annual review of cell and developmental biology* **19**, 237-259, doi:10.1146/annurev.cellbio.19.111401.092717 (2003).

61 Heard, E. & Disteche, C. M. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes & development* **20**, 1848-1867, doi:10.1101/gad.1422906 (2006).

62 Lelli, K. M., Slattery, M. & Mann, R. S. Disentangling the many layers of eukaryotic transcriptional regulation. *Annual review of genetics* **46**, 43-68, doi:10.1146/annurev-genet-110711-155437 (2012).

63 Coulon, A., Chow, C. C., Singer, R. H. & Larson, D. R. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature reviews. Genetics* **14**, 572-584, doi:10.1038/nrg3484 (2013).

64 Lickwar, C. R., Mueller, F., Hanlon, S. E., McNally, J. G. & Lieb, J. D. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* **484**, 251-255, doi:10.1038/nature10985 (2012).

65 Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461**, 186-192, doi:10.1038/nature08449 (2009).

66 Orphanides, G., Lagrange, T. & Reinberg, D. The general transcription factors of RNA polymerase II. *Genes & development* **10**, 2657-2683 (1996).

67 Roeder, R. G., Schwartz, L. B. & Sklar, V. E. Function, structure, and regulation of eukaryotic nuclear RNA polymerases. *The symposium / The Society for Developmental Biology. Society for Developmental Biology. Symposium*, 29-52 (1976).

68 Matsui, T., Segall, J., Weil, P. A. & Roeder, R. G. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *The Journal of biological chemistry* **255**, 11992-11996 (1980).

69 Sawadogo, M. & Roeder, R. G. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 4394-4398 (1985).

70 Roeder, R. G. The role of general initiation factors in transcription by RNA polymerase II. *Trends in biochemical sciences* **21**, 327-335 (1996).

71 Gangloff, Y. G., Romier, C., Thuault, S., Werten, S. & Davidson, I. The histone fold is a key structural motif of transcription factor TFIID. *Trends in biochemical sciences* **26**, 250-257 (2001).

72 Malkowska, M., Kokoszynska, K., Rychlewski, L. & Wyrwicz, L. Structural bioinformatics of the general transcription factor TFIID. *Biochimie* **95**, 680-691, doi:10.1016/j.biochi.2012.10.024 (2013).

73 Thomas, M. C. & Chiang, C. M. The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology* **41**, 105-178, doi:10.1080/10409230600648736 (2006).

74 Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58-69, doi:10.1016/j.cell.2007.08.016 (2007).

75 Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-1425 (2000).

76 Bieniossek, C. *et al.* The architecture of human general transcription factor TFIID core complex. *Nature* **493**, 699-702, doi:10.1038/nature11791 (2013).

77 Trowitzsch, S. *et al.* Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nature communications* **6**, 6011, doi:10.1038/ncomms7011 (2015).

78 Guermah, M., Ge, K., Chiang, C. M. & Roeder, R. G. The TBN protein, which is essential for early embryonic mouse development, is an inducible TAFII

implicated in adipogenesis. *Molecular cell* **12**, 991-1001 (2003).

79 Tatarakis, A. *et al.* Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Molecular cell* **31**, 531-543, doi:10.1016/j.molcel.2008.07.013 (2008).

80 Pijnappel, W. W. *et al.* A central role for TFIID in the pluripotent transcription circuitry. *Nature* **495**, 516-519, doi:10.1038/nature11970 (2013).

81 Williamson, P. R. & Kagan, H. M. Reaction pathway of bovine aortic lysyl oxidase. *The Journal of biological chemistry* **261**, 9477-9482 (1986).

82 Siegel, R. C. Biosynthesis of collagen crosslinks: increased activity of purified lysyl oxidase with reconstituted collagen fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 4826-4830 (1974).

83 Kuivaniemi, H., Korhonen, R. M., Vaheri, A. & Kivirikko, K. I. Deficient production of lysyl oxidase in cultures of malignantly transformed human cells. *FEBS letters* **195**, 261-264 (1986).

84 Lazarus, H. M., Cruikshank, W. W., Narasimhan, N., Kagan, H. M. & Center, D. M. Induction of human monocyte motility by lysyl oxidase. *Matrix biology : journal of the International Society for Matrix Biology* **14**, 727-731 (1995).

85 Li, W., Liu, G., Chou, I. N. & Kagan, H. M. Hydrogen peroxide-mediated, lysyl oxidase-dependent chemotaxis of vascular smooth muscle cells. *Journal of cellular biochemistry* **78**, 550-557 (2000).

86 Erler, J. T. *et al.* Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* **440**, 1222-1226, doi:10.1038/nature04695 (2006).

87 Barker, H. E., Cox, T. R. & Erler, J. T. The rationale for targeting the LOX family in cancer. *Nature reviews. Cancer* **12**, 540-552, doi:10.1038/nrc3319 (2012).

88 Mello, M. L., Alvarenga, E. M., Vidal Bde, C. & Di Donato, A. Chromatin supraorganization, mitotic abnormalities and proliferation in cells with increased or down-regulated lox expression: Indirect evidence of a LOX-histone H1 interaction in vivo. *Micron* **42**, 8-16, doi:10.1016/j.micron.2010.09.001 (2011).

89 Oleggini, R. & Di Donato, A. Lysyl oxidase regulates MMTV promoter: indirect evidence of histone H1 involvement. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **89**, 522-532, doi:10.1139/o11-049 (2011).

90 Peinado, H. *et al.* A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression. *The EMBO journal* **24**, 3446-3458, doi:10.1038/sj.emboj.7600781 (2005).

91 Herranz, N. *et al.* Lysyl oxidase-like 2 deaminates lysine 4 in histone H3. *Molecular cell* **46**, 369-376, doi:10.1016/j.molcel.2012.03.002 (2012).

92 Millanes-Romero, A. *et al.* Regulation of heterochromatin transcription by Snail1/LOXL2 during epithelial-to-mesenchymal transition. *Molecular cell* **52**, 746-757, doi:10.1016/j.molcel.2013.10.015 (2013).

93 Martin, A. *et al.* Lysyl oxidase-like 2 represses Notch1 expression in the skin to promote squamous cell

carcinoma progression. *The EMBO journal* **34**, 1090-1109, doi:10.15252/embj.201489975 (2015).

94 Barry-Hamilton, V. *et al.* Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathologic microenvironment. *Nature medicine* **16**, 1009-1017, doi:10.1038/nm.2208 (2010).

95 Peng, L. *et al.* Secreted LOXL2 is a novel therapeutic target that promotes gastric cancer metastasis via the Src/FAK pathway. *Carcinogenesis* **30**, 1660-1669, doi:10.1093/carcin/bgp178 (2009).

96 Barker, H. E. *et al.* LOXL2-mediated matrix remodeling in metastasis and mammary gland involution. *Cancer research* **71**, 1561-1572, doi:10.1158/0008-5472.CAN-10-2868 (2011).

97 Wong, C. C. *et al.* Lysyl oxidase-like 2 is critical to tumor microenvironment and metastatic niche formation in hepatocellular carcinoma. *Hepatology* **60**, 1645-1658, doi:10.1002/hep.27320 (2014).

98 Lim, J. & Thiery, J. P. Epithelial-mesenchymal transitions: insights from development. *Development* **139**, 3471-3486, doi:10.1242/dev.071209 (2012).

99 Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation* **119**, 1420-1428, doi:10.1172/JCI39104 (2009).

100 Scheel, C. & Weinberg, R. A. Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Seminars in cancer biology* **22**, 396-403, doi:10.1016/j.semcancer.2012.04.001 (2012).

101 Zeisberg, M. & Kalluri, R. Cellular mechanisms of tissue fibrosis. 1. Common and organ-specific mechanisms associated with tissue fibrosis. *American journal of physiology. Cell physiology* **304**, C216-225, doi:10.1152/ajpcell.00328.2012 (2013).

102 Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nature reviews. Molecular cell biology* **15**, 178-196, doi:10.1038/nrm3758 (2014).

103 De Craene, B. & Berx, G. Regulatory networks defining EMT during cancer initiation and progression. *Nature reviews. Cancer* **13**, 97-110, doi:10.1038/nrc3447 (2013).

104 Mendez, M. G., Kojima, S. & Goldman, R. D. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **24**, 1838-1851, doi:10.1096/fj.09-151639 (2010).

105 Peinado, H., Olmeda, D. & Cano, A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nature reviews. Cancer* **7**, 415-428, doi:10.1038/nrc2131 (2007).

106 Eastham, A. M. *et al.* Epithelial-mesenchymal transition events during human embryonic stem cell differentiation. *Cancer research* **67**, 11254-11262, doi:10.1158/0008-5472.CAN-07-2253 (2007).

107 Acloque, H., Adams, M. S., Fishwick, K., Bronner-Fraser, M. & Nieto, M. A. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *The Journal of clinical investigation* **119**, 1438-1449, doi:10.1172/JCI38019 (2009).

108 Esteban, M. A. *et al.* The mesenchymal-to-epithelial transition in somatic cell reprogramming. *Current opinion in genetics & development* **22**, 423-428, doi:10.1016/j.gde.2012.09.004 (2012).

109 Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704-715, doi:10.1016/j.cell.2008.03.027 (2008).

110 Brabletz, T., Jung, A., Spaderna, S., Hlubek, F. & Kirchner, T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nature reviews. Cancer* **5**, 744-749, doi:10.1038/nrc1694 (2005).

111 Schietke, R. *et al.* The lysyl oxidases LOX and LOXL2 are necessary and sufficient to repress E-cadherin in hypoxia: insights into cellular transformation processes mediated by HIF-1. *The Journal of biological chemistry* **285**, 6658-6669, doi:10.1074/jbc.M109.042424 (2010).

112 Voloshenyuk, T. G., Landesman, E. S., Khoutorova, E., Hart, A. D. & Gardner, J. D. Induction of cardiac fibroblast lysyl oxidase by TGF-beta1 requires PI3K/Akt, Smad3, and MAPK signaling. *Cytokine* **55**, 90-97, doi:10.1016/j.cyto.2011.03.024 (2011).

113 Fujimoto, E. & Tajima, S. Reciprocal regulation of LOX and LOXL2 expression during cell adhesion and terminal differentiation in epidermal keratinocytes. *Journal of dermatological science* **55**, 91-98, doi:10.1016/j.jdermsci.2009.03.010 (2009).

114 Slaga, T. J., Budunova, I. V., Gimenez-Conti, I. B. & Aldaz, C. M. The mouse skin carcinogenesis model. *The journal of investigative dermatology. Symposium proceedings / the Society for Investigative Dermatology, Inc. [and] European Society for Dermatological Research* **1**, 151-156 (1996).

115 Kim, J. H. *et al.* The role of lysyl oxidase-like 2 in the odontogenic differentiation of human dental pulp stem cells. *Molecules and cells* **35**, 543-549, doi:10.1007/s10059-013-0080-3 (2013).

116 Iftikhar, M. *et al.* Lysyl oxidase-like-2 (LOXL2) is a major isoform in chondrocytes and is critically required for differentiation. *The Journal of biological chemistry* **286**, 909-918, doi:10.1074/jbc.M110.155622 (2011).

117 Goodell, M. A., Nguyen, H. & Shroyer, N. Somatic stem cell heterogeneity: diversity in the blood, skin and intestinal stem cell compartments. *Nature reviews. Molecular cell biology* **16**, 299-309, doi:10.1038/nrm3980 (2015).

118 Acloque, H., Ocana, O. H. & Nieto, M. A. Mutual exclusion of transcription factors and cell behaviour in the definition of vertebrate embryonic territories. *Current opinion in genetics & development* **22**, 308-314, doi:10.1016/j.gde.2012.03.004 (2012).

119 Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871-890, doi:10.1016/j.cell.2009.11.007 (2009).

120 Niakan, K. K., Schrode, N., Cho, L. T. & Hadjantonakis, A. K. Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. *Nature protocols* **8**, 1028-1041, doi:10.1038/nprot.2013.049 (2013).

121 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).

122 Brook, F. A. & Gardner, R. L. The origin and efficient derivation of embryonic stem cells in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 5709-5712 (1997).

123 Williams, R. L. *et al.* Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687, doi:10.1038/336684a0 (1988).

124 Smith, A. G. *et al.* Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-690, doi:10.1038/336688a0 (1988).

125 Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523, doi:10.1038/nature06968 (2008).

126 Tam, P. P. & Rossant, J. Mouse embryonic chimeras: tools for studying mammalian development. *Development* **130**, 6155-6163, doi:10.1242/dev.00893 (2003).

127 Niwa, H. Mouse ES cell culture system as a model of development. *Development, growth & differentiation* **52**, 275-283, doi:10.1111/j.1440-169X.2009.01166.x (2010).

128 Okamoto, K. *et al.* A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**, 461-472 (1990).

129 Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391 (1998).

130 Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation,

dedifferentiation or self-renewal of ES cells. *Nature genetics* **24**, 372-376, doi:10.1038/74199 (2000).

131 Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nature cell biology* **9**, 625-635, doi:10.1038/ncb1589 (2007).

132 Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106-1117, doi:10.1016/j.cell.2008.04.043 (2008).

133 Keramari, M. *et al.* Sox2 is essential for formation of trophectoderm in the preimplantation embryo. *PloS one* **5**, e13952, doi:10.1371/journal.pone.0013952 (2010).

134 Kopp, J. L., Ormsbee, B. D., Desler, M. & Rizzino, A. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem cells* **26**, 903-911, doi:10.1634/stemcells.2007-0951 (2008).

135 Plusa, B., Piliszek, A., Frankenberg, S., Artus, J. & Hadjantonakis, A. K. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**, 3081-3091, doi:10.1242/dev.021519 (2008).

136 Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-737, doi:10.1016/j.cell.2009.07.039 (2009).

137 Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-1234, doi:10.1038/nature06403 (2007).

138 Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655 (2003).

139 Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics* **38**, 431-440, doi:10.1038/ng1760 (2006).

140 Nichols, J. & Smith, A. Pluripotency in the embryo and in culture. *Cold Spring Harbor perspectives in biology* **4**, a008128, doi:10.1101/cshperspect.a008128 (2012).

141 Martello, G. & Smith, A. The nature of embryonic stem cells. *Annual review of cell and developmental biology* **30**, 647-675, doi:10.1146/annurev-cellbio-100913-013116 (2014).

142 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).

143 Tsubooka, N. *et al.* Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes to cells : devoted to molecular & cellular mechanisms* **14**, 683-694, doi:10.1111/j.1365-2443.2009.01301.x (2009).

144 Tam, P. P. & Loebel, D. A. Gene function in mouse embryogenesis: get set for gastrulation. *Nature reviews. Genetics* **8**, 368-381, doi:10.1038/nrg2084 (2007).

145 Johnson, M. H. From mouse egg to mouse embryo: polarities, axes, and tissues. *Annual review of cell and developmental biology* **25**, 483-512, doi:10.1146/annurev.cellbio.042308.113348 (2009).

146 Kalmar, T. *et al.* Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS biology* **7**, e1000149, doi:10.1371/journal.pbio.1000149 (2009).

147 Hayashi, K., Lopes, S. M., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell stem cell* **3**, 391-401, doi:10.1016/j.stem.2008.07.027 (2008).

148 Kalkan, T. & Smith, A. Mapping the route from naive pluripotency to lineage specification. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **369**, doi:10.1098/rstb.2013.0540 (2014).

149 Betschinger, J. *et al.* Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335-347, doi:10.1016/j.cell.2013.03.012 (2013).

150 Loh, K. M. & Lim, B. A precarious balance: pluripotency factors as lineage specifiers. *Cell stem cell* **8**, 363-369, doi:10.1016/j.stem.2011.03.013 (2011).

151 Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532, doi:10.1016/j.cell.2011.06.052 (2011).

152 Ozair, M. Z., Kintner, C. & Brivanlou, A. H. Neural induction and early patterning in vertebrates. *Wiley interdisciplinary reviews. Developmental biology* **2**, 479-498, doi:10.1002/wdev.90 (2013).

153 Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature biotechnology* **27**, 275-280, doi:10.1038/nbt.1529 (2009).

154 Chng, Z., Teo, A., Pedersen, R. A. & Vallier, L. SIP1 mediates cell-fate decisions between neuroectoderm and mesendoderm in human pluripotent stem cells. *Cell stem cell* **6**, 59-70, doi:10.1016/j.stem.2009.11.015 (2010).

155 Kamiya, D. *et al.* Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* **470**, 503-509, doi:10.1038/nature09726 (2011).

156 Moody, S. A., Klein, S. L., Karpinski, B. A., Maynard, T. M. & Lamantia, A. S. On becoming neural: what the embryo can tell us about differentiating neural stem cells. *American journal of stem cells* **2**, 74-94 (2013).

157 Maden, M. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nature reviews. Neuroscience* **8**, 755-765, doi:10.1038/nrn2212 (2007).

158 Dziewczapolski, G., Lie, D. C., Ray, J., Gage, F. H. & Shults, C. W. Survival and differentiation of adult rat-derived neural progenitor cells transplanted to the striatum of hemiparkinsonian rats. *Experimental neurology* **183**, 653-664 (2003).

159 Xu, J. *et al.* Retinoic acid promotes neural conversion of mouse embryonic stem cells in adherent monoculture. *Molecular biology reports* **39**, 789-795, doi:10.1007/s11033-011-0800-8 (2012).

160 Blomhoff, R. & Blomhoff, H. K. Overview of retinoid metabolism and function. *Journal of neurobiology* **66**, 606-630, doi:10.1002/neu.20242 (2006).

161 Kawaguchi, R. *et al.* A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A. *Science* **315**, 820-825, doi:10.1126/science.1136244 (2007).

162 Sandell, L. L. *et al.* RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. *Genes & development* **21**, 1113-1124, doi:10.1101/gad.1533407 (2007).

163 Budhu, A. S. & Noy, N. Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and cellular biology* **22**, 2632-2641 (2002).

164 Bastien, J. & Rochette-Egly, C. Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene* **328**, 1-16, doi:10.1016/j.gene.2003.12.005 (2004).

165 Ross, A. C. & Zolfaghari, R. Cytochrome P450s in the regulation of cellular retinoic acid metabolism. *Annual review of nutrition* **31**, 65-87, doi:10.1146/annurev-nutr-072610-145127 (2011).

166 Hart, S. M. Modulation of nuclear receptor dependent transcription. *Biological research* **35**, 295-303 (2002).

167 Wei, L. N. Retinoid receptors and their coregulators. *Annual review of pharmacology and toxicology* **43**, 47-72, doi:10.1146/annurev.pharmtox.43.100901.140301 (2003).

168 Cunningham, T. J. & Duester, G. Mechanisms of retinoic acid signalling and its roles in organ and limb development. *Nature reviews. Molecular cell biology* **16**, 110-123, doi:10.1038/nrm3932 (2015).

169 Soprano, D. R., Teets, B. W. & Soprano, K. J. Role of retinoic acid in the differentiation of embryonal carcinoma and embryonic stem cells. *Vitamins and hormones* **75**, 69-95, doi:10.1016/S0083-6729(06)75003-8 (2007).

170 Mahony, S. *et al.* Ligand-dependent dynamics of retinoic acid receptor binding during early neurogenesis. *Genome biology* **12**, R2, doi:10.1186/gb-2011-12-1-r2 (2011).

171 Shahhoseini, M., Taei, A., Mehrjardi, N. Z., Salekdeh, G. H. & Baharvand, H. Epigenetic analysis of human embryonic carcinoma cells during retinoic acid-induced neural differentiation. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **88**, 527-538, doi:10.1139/o09-181 (2010).

172 Rosa, A. & Brivanlou, A. H. A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *The EMBO journal* **30**, 237-248, doi:10.1038/emboj.2010.319 (2011).

173 Martinez-Ceballos, E. & Gudas, L. J. Hoxa1 is required for the retinoic acid-induced differentiation of embryonic stem cells into neurons. *Journal of neuroscience research* **86**, 2809-2819, doi:10.1002/jnr.21729 (2008).

174 Gu, P. *et al.* Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell

differentiation. *Molecular and cellular biology* **25**, 8507-8519, doi:10.1128/MCB.25.19.8507-8519.2005 (2005).

175 Fuhrmann, G. *et al.* Mouse germline restriction of Oct4 expression by germ cell nuclear factor. *Developmental cell* **1**, 377-387 (2001).

176 Zhang, X., Wen, H. & Shi, X. Lysine methylation: beyond histones. *Acta biochimica et biophysica Sinica* **44**, 14-27, doi:10.1093/abbs/gmr100 (2012).

177 Kouskouti, A., Scheer, E., Staub, A., Tora, L. & Talianidis, I. Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Molecular cell* **14**, 175-182 (2004).

178 Nishioka, K. *et al.* Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes & development* **16**, 479-489, doi:10.1101/gad.967202 (2002).

179 Couture, J. F., Collazo, E., Hauk, G. & Trievel, R. C. Structural basis for the methylation site specificity of SET7/9. *Nature structural & molecular biology* **13**, 140-146, doi:10.1038/nsmb1045 (2006).

180 Soutoglou, E. *et al.* The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Molecular and cellular biology* **25**, 4092-4104, doi:10.1128/MCB.25.10.4092-4104.2005 (2005).

181 Mohan, W. S., Jr., Scheer, E., Wendling, O., Metzger, D. & Tora, L. TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Molecular and cellular biology* **23**, 4307-4318 (2003).

182 Farah, M. H. *et al.* Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development* **127**, 693-702 (2000).

183 Tabin, C. J. Retinoids, homeoboxes, and growth factors: toward molecular models for limb development. *Cell* **66**, 199-217 (1991).

184 De Kumar, B. *et al.* Analysis of dynamic changes in retinoid-induced transcription and epigenetic profiles of murine Hox clusters in ES cells. *Genome research* **25**, 1229-1243, doi:10.1101/gr.184978.114 (2015).

185 Gansner, J. M., Mendelsohn, B. A., Hultman, K. A., Johnson, S. L. & Gitlin, J. D. Essential role of lysyl oxidases in notochord development. *Developmental biology* **307**, 202-213, doi:10.1016/j.ydbio.2007.04.029 (2007).

186 Rizzino, A. Concise review: The Sox2-Oct4 connection: critical players in a much larger interdependent network integrated at multiple levels. *Stem cells* **31**, 1033-1039, doi:10.1002/stem.1352 (2013).

187 Bylund, M., Andersson, E., Novitch, B. G. & Muhr, J. Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nature neuroscience* **6**, 1162-1168, doi:10.1038/nn1131 (2003).

188 Huang, J. *et al.* p53 is regulated by the lysine demethylase LSD1. *Nature* **449**, 105-108, doi:10.1038/nature06092 (2007).

189 Pless, O. *et al.* G9a-mediated lysine methylation alters the function of CCAAT/enhancer-binding protein-beta. *The Journal of biological chemistry* **283**, 26357-26363, doi:10.1074/jbc.M802132200 (2008).

190 Zhang, K. *et al.* The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* **122**, 723-734, doi:10.1016/j.cell.2005.06.021 (2005).

191 Gingras, A. C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nature reviews. Molecular cell biology* **8**, 645-654, doi:10.1038/nrm2208 (2007).

192 Csiszar, K. Lysyl oxidases: a novel multifunctional amine oxidase family. *Progress in nucleic acid research and molecular biology* **70**, 1-32 (2001).

193 Kim, Y. M., Kim, E. C. & Kim, Y. The human lysyl oxidase-like 2 protein functions as an amine oxidase toward collagen and elastin. *Molecular biology reports* **38**, 145-149, doi:10.1007/s11033-010-0088-0 (2011).

194 Wisniewski, J. R., Zougman, A., Kruger, S. & Mann, M. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Molecular & cellular proteomics : MCP* **6**, 72-87, doi:10.1074/mcp.M600255-MCP200 (2007).

195 Daujat, S., Zeissler, U., Waldmann, T., Happel, N. & Schneider, R. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *The Journal of biological chemistry* **280**, 38090-38095, doi:10.1074/jbc.C500229200 (2005).

196 Maison, C. *et al.* SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nature genetics* **43**, 220-227, doi:10.1038/ng.765 (2011).

197 Kagan, H. M., Williams, M. A., Calaman, S. D. & Berkowitz, E. M. Histone H1 is a substrate for lysyl oxidase and contains endogenous sodium borotritide-reducible residues. *Biochemical and biophysical research communications* **115**, 186-192 (1983).

198 Giampuzzi, M., Oleggini, R. & Di Donato, A. Demonstration of in vitro interaction between tumor suppressor lysyl oxidase and histones H1 and H2: definition of the regions involved. *Biochimica et biophysica acta* **1647**, 245-251 (2003).

199 Oleggini, R., Gastaldo, N. & Di Donato, A. Regulation of elastin promoter by lysyl oxidase and growth factors: cross control of lysyl oxidase on TGF-beta1 effects. *Matrix biology : journal of the International Society for Matrix Biology* **26**, 494-505, doi:10.1016/j.matbio.2007.02.003 (2007).

200 Giampuzzi, M. *et al.* Lysyl oxidase activates the transcription activity of human collagene III promoter. Possible involvement of Ku antigen. *The Journal of biological chemistry* **275**, 36341-36349, doi:10.1074/jbc.M003362200 (2000).

201 Giampuzzi, M., Oleggini, R., Albanese, C., Pestell, R. & Di Donato, A. beta-catenin signaling and regulation of cyclin D1 promoter in NRK-49F cells transformed by down-regulation of the tumor suppressor lysyl oxidase. *Biochimica et biophysica acta* **1745**, 370-381, doi:10.1016/j.bbamcr.2005.04.012 (2005).

202 Roque, A., Ponte, I., Arrondo, J. L. & Suau, P. Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. *Nucleic acids research* **36**, 4719-4726, doi:10.1093/nar/gkn440 (2008).

203 Haushalter, K. A. & Kadonaga, J. T. Chromatin assembly by DNA-translocating motors. *Nature reviews. Molecular cell biology* **4**, 613-620, doi:10.1038/nrm1177 (2003).

204 Quivy, J. P. *et al.* A CAF-1 dependent pool of HP1 during heterochromatin duplication. *The EMBO journal* **23**, 3516-3526, doi:10.1038/sj.emboj.7600362 (2004).

205 Murzina, N., Verreault, A., Laue, E. & Stillman, B. Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. *Molecular cell* **4**, 529-540 (1999).

206 Yamamoto, K. & Sonoda, M. Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1. *Biochemical and biophysical research communications* **301**, 287-292 (2003).

207 Stracker, T. H. & Petrini, J. H. The MRE11 complex: starting from the ends. *Nature reviews. Molecular cell biology* **12**, 90-103, doi:10.1038/nrm3047 (2011).

208 Cao, X. J., Arnaudo, A. M. & Garcia, B. A. Large-scale global identification of protein lysine methylation in vivo. *Epigenetics* **8**, 477-485, doi:10.4161/epi.24547 (2013).

209 Kurash, J. K. *et al.* Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. *Molecular cell* **29**, 392-400, doi:10.1016/j.molcel.2007.12.025 (2008).

210 Munro, S., Khaire, N., Inche, A., Carr, S. & La Thangue, N. B. Lysine methylation regulates the pRb tumour suppressor protein. *Oncogene* **29**, 2357-2367, doi:10.1038/onc.2009.511 (2010).

211 Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-1558, doi:10.1126/science.1235122 (2013).

212 Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319**, 1352-1355, doi:10.1126/science.1140735 (2008).

213 Burrell, R. A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338-345, doi:10.1038/nature12625 (2013).

214 Barber, L. J. *et al.* Secondary mutations in BRCA2 associated with clinical resistance to a PARP inhibitor. *The Journal of pathology* **229**, 422-429, doi:10.1002/path.4140 (2013).

215 Mereniuk, T. R. *et al.* Synthetic lethal targeting of PTEN-deficient cancer cells using selective disruption of polynucleotide kinase/phosphatase. *Molecular cancer therapeutics* **12**, 2135-2144, doi:10.1158/1535-7163.MCT-12-1093 (2013).

216 Pearl, L. H., Schierz, A. C., Ward, S. E., Al-Lazikani, B. & Pearl, F. M. Therapeutic opportunities within the DNA damage response. *Nature reviews. Cancer* **15**, 166-180, doi:10.1038/nrc3891 (2015).

217 Riganti, C., Gazzano, E., Polimeni, M., Aldieri, E. & Ghigo, D. The pentose phosphate pathway: an antioxidant defense and a crossroad in tumor cell fate. *Free radical biology & medicine* **53**, 421-436, doi:10.1016/j.freeradbiomed.2012.05.006 (2012).

218 Boros, L. G. *et al.* Inhibition of the oxidative and nonoxidative pentose phosphate pathways by

somatostatin: a possible mechanism of antitumor action. *Medical hypotheses* **50**, 501-506 (1998).

219 Jiang, P. *et al.* p53 regulates biosynthesis through direct inactivation of glucose-6-phosphate dehydrogenase. *Nature cell biology* **13**, 310-316, doi:10.1038/ncb2172 (2011).

220 Ahmad, A. *et al.* Phosphoglucose isomerase/autocrine motility factor mediates epithelial-mesenchymal transition regulated by miR-200 in breast cancer cells. *Cancer research* **71**, 3400-3409, doi:10.1158/0008-5472.CAN-10-0965 (2011).

221 Grunberg, S. & Hahn, S. Structural insights into transcription initiation by RNA polymerase II. *Trends in biochemical sciences* **38**, 603-611, doi:10.1016/j.tibs.2013.09.002 (2013).

222 Samara, N. L. & Wolberger, C. A new chapter in the transcription SAGA. *Current opinion in structural biology* **21**, 767-774, doi:10.1016/j.sbi.2011.09.004 (2011).

223 Metzger, D., Scheer, E., Soldatov, A. & Tora, L. Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. *The EMBO journal* **18**, 4823-4834, doi:10.1093/emboj/18.17.4823 (1999).

224 Huq, M. D., Tsai, N. P., Khan, S. A. & Wei, L. N. Lysine trimethylation of retinoic acid receptor-alpha: a novel means to regulate receptor function. *Molecular & cellular proteomics : MCP* **6**, 677-688, doi:10.1074/mcp.M600223-MCP200 (2007).

225 Reynaud, C., Baas, D., Gleyzal, C., Le Guellec, D. & Sommer, P. Morpholino knockdown of lysyl oxidase

impairs zebrafish development, and reflects some aspects of copper metabolism disorders. *Matrix biology : journal of the International Society for Matrix Biology* **27**, 547-560, doi:10.1016/j.matbio.2008.03.002 (2008).

226 Chen, C., Nott, T. J., Jin, J. & Pawson, T. Deciphering arginine methylation: Tudor tells the tale. *Nature reviews. Molecular cell biology* **12**, 629-642, doi:10.1038/nrm3185 (2011).

227 Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953, doi:10.1016/j.cell.2004.12.012 (2004).

228 Shi, Y. & Whetstine, J. R. Dynamic regulation of histone lysine methylation by demethylases. *Molecular cell* **25**, 1-14, doi:10.1016/j.molcel.2006.12.010 (2007).

229 Katoh, M. & Katoh, M. Identification and characterization of JMJD2 family genes in silico. *International journal of oncology* **24**, 1623-1628 (2004).

230 Cuthbert, G. L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545-553, doi:10.1016/j.cell.2004.08.020 (2004).

231 Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306**, 279-283, doi:10.1126/science.1101400 (2004).

232 Nakashima, K., Hagiwara, T. & Yamada, M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *The Journal of biological chemistry* **277**, 49562-49568, doi:10.1074/jbc.M208795200 (2002).

233 Jones, J. E., Causey, C. P., Knuckley, B., Slack-Noyes, J. L. & Thompson, P. R. Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Current opinion in drug discovery & development* **12**, 616-627 (2009).

234 Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* **2**, 1896-1906, doi:10.1038/nprot.2007.261 (2007).

235 Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567, doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2 (1999).

236 Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic acids research* **31**, e15 (2003).

237 Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* **3**, Article3, doi:10.2202/1544-6115.1027 (2004).

238 Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **203**, 253-310, doi:10.1002/aja.1002030302 (1995).

239 Streit, A. *et al.* Preventing the loss of competence for neural induction: HGF/SF, L5 and Sox-2. *Development* **124**, 1191-1202 (1997).

240 Blader, P., Fischer, N., Gradwohl, G., Guillemot, F. & Strahle, U. The activity of neurogenin1 is controlled by local cues in the zebrafish embryo. *Development* **124**, 4557-4569 (1997).

241 Thisse, B. *et al.* Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods in cell biology* **77**, 505-519 (2004).

RESEARCH ARTICLES

Research articles resulting from this thesis:

Iturbide A, Pascual-Reguant L, Fargas L, Cebriá JP, Alsina B, García de Herreros A, Peiró S. LOXL2 Oxidizes Methylated TAF10 and Controls TFIID-Dependent Genes during Neural Progenitor Differentiation. Molecular Cell, 2015, 58(5): 755-766.

Accession numbers: GSE65675

Collaboration with a project to determine the role of LOXL2 together with Snail1 in the regulation of heterochromatin transcription during EMT:

Millanes-Romero A, Herranz N, Perrera V, **Iturbide A**, Loubat J, Jenuwein T, García de Herreros A, Peiró S. Regulation of Heterochromatin Transcription by Snail1/LOXL2 During Epithelial to Mesenchymal Transition. Molecular Cell, 2013, 52(5): 746-757.

Review articles:

Iturbide A, García de Herreros A, Peiró S. A New Role for LOX and LOXL2 Proteins in Transcription Regulation. *FEBS J*, 2015, 282(9): 1768-1773.

Iturbide A, Pascual-Reguant L, Fargas L, Cebrià JP, Alsina B, García de Herreros A, Peiró S. LOXL2 Oxidizes Methylated TAF10 and Controls TFIID-Dependent Genes during Neural Progenitor Differentiation. Mol Cell. 2015 Jun 4;58(5):755-66. doi: 10.1016/j.molcel.2015.04.012

ACKNOWLEDGEMENTS

Gauzak ez dira beti batek espero dituen bezala suertatzen. Orain dela hiru urte gaur nagoen tokira iritsiko nintzela norbaitek esan izan balit, ez nukeen hitz bat ere sinistuko. Eta hemen nago, amets bat iadanik beteta eta beste asko betetzeko asmoz. Ezin ahaztu bide honetan ez naizela bakarrik izan, asko izan dira nirekin ibili dutenak. Batzuk denbora gehiagoz, besteek gutxiagoz, baina denek lagundu naute helmugara iristen, baita oraindik ibili beharko ditudan bideak erakusten ere. Zuei guztiei, eskerrak emain nahi dizkizuet.

En primer lugar te quiero dar las gracias a ti Sandra por darme la oportunidad de realizar esta tesis. No solo has sabido enseñarme lo necesario para sacar un proyecto adelante, sino a ser más crítica, a pensar y repensar las cosas, a plantearme preguntas y a esperar siempre más. Pero, por encima de todo, me has transmitido tu pasión por la ciencia. No me imagino haciendo otra cosa, y eso es seguro en gran parte gracias a ti. Sé que en ocasiones no es fácil trabajar conmigo, puedo llegar a ser muy cabezota y pesimista, y seguro que te he dado más de un dolor de cabeza. Gracias por seguir ahí a pesar de todo. También quiero agradecer su ayuda y buenos consejos a Antonio, Jepi y Victor. A pesar de no ser responsabilidad vuestra siempre habéis estado ahí para echarme una mano cuando hiciera falta, gracias por todo.

No puedo olvidarme de todas las personas que me acogieron en el laboratorio cuando llegué: Silvia, Nuria, Alba, Estel, Jou y Ali, gracias por ayudarme cuando era una estudiante desastrosa que no sabía ni preparar un buffer. Jelena, thank you for saving me so many times in the lab. You are a great scientist and I have no doubt that you will succed. Rosa, puedo decir sin miedo a equivocarme que si no fuera por ti no habría llegado a donde estoy hoy. Gracias por todas las broncas que me metiste cuando era estudiante, por las correcciones, por enseñarme a hacer bien las cosas, por pensar que merecía la pena invertir ese tiempo en mi. Pero sobre todo, gracias por ser tan buena amiga, siempre he podido contar contigo y eso ha marcado una gran diferencia.

Mis primeras aventuras por el lab no habrían sido las mismas sin mis compañeros de master, Josu y Estefanía. Fani, sin tu buena energía y sin tus chistes malos por minuto el lab nunca ha vuelto a ser el mismo. Josu, zikina, sobran palabras. Bien sabes que siempre nos hemos apoyado, en los buenos y malos momentos, y hemos tirado el uno del otro para seguir adelante. Es un respiro poder contar con un amigo tan bueno. Sigue trabajando así de duro, estoy segura de que todo saldrá bien.

Y a todos los que han compartido laboratorio conmigo hasta el final de mi tesis, también os quiero dar las gracias. Rocco, Lorena y Alex, os merecéis un premio solo por haberme

tenido que aguantar tanto tiempo, gracias por el buen ambiente sobre todo en el despacho en esta recta final de la tesis, cuando tanto hace falta. Alex, creo que me has ayudado más de lo que eres consciente. Tu buena energía, tu fortaleza y tu optimismo siempre me han inspirado a mirar las cosas con más filosofía. Haces que los malos momentos no lo sean tanto y que la gente a tu alrededor se sienta mejor, eskerrik asko por ser así. Raul, tu sí que has tenido que aguantar a la vasca de principio a fin. Gracias por las muchas veces que me has ayudado y por cuidar de todos nosotros.

Gracias también a Chromatin team por todos los momentos pasados juntos. Jess, me ha encantado poder descubrirte tras la chica tímida que venía lab y apenas saludaba. Has florecido de la mejor manera posible, no cambies y sigue con ese buen humor y generosidad que te caracterizan. Joan Pau, cuando llegaste al grupo te tocó ser mi estudiante cuando yo apenas sabía hacer un western. Fuimos un desastre, pero lo fuimos juntos. Hemos aprendido muchas cosas de la mano y pasado muy buenos momentos, gracias por los buenos recuerdos. Laura, a ti también te tocó aguantarme, aunque quiero pensar que en tu caso aprendiste más cosas conmigo. Además de las horas de trabajo, hemos compartido proyecto y la alegría de trabajar con las ESCs. Gracias por trabajar tan duro y ayudarme a sacar el paper adelante. Gemma, la última incorporación a

los chromatin! No solo tienes un gran sentido de la responsabilidad sino también una fantástica actitud para el trabajo. Sigue así, no tengo ninguna duda de que superarás todos los obstáculos y harás una gran tesis. También quiero acordarme de las últimas incorporaciones al Snail team: Willy, Laura, Rubentxo, Aida y María, por venir con tan buena energía. Estoy segura de que los snailitos seguirán manteniendo su esencia con esta nueva generación, además de publicar varios natures of course.

La tesis no habría sido la misma sin los buenos vecinos que he tenido. Quiero agradecer a los Bigas que me dejaran robar material de western y anticuerpos siempre que me hizo falta. Además del buen ambiente que aportan a la galería, los chistes cruzados entre labs y por ofrecernos siempre comida. Bereziki zuri eman nahi dizkizut eskerrak Erika, ez dakit zenbatetan eskatu dizudan laguntza, eta lanpetuta ibili arren, beti aurkitu duzu niretzako momentu bat, milesker. También quiero acordarme de las chicas de “al lado’s lab”. Conchi, mi compañera en la galería, muchas gracias por los buenos ratos y por tus “zascas” de campeonato. Silvia, no he conocido a nadie con tan buena energía, siempre con una sonrisa en la cara, me has alegrado muchos días. Gracias por transmitir ese buen humor, por cargarnos las pilas y por dejarme siempre el último sashimi de salmón. Neus, no eres solo la referencia en tu lab sino muchas veces también en el nuestro, gracias por todo.

También quiero dar las gracias a mi querido compañero en quejas, lamentos, frustraciones, ranciedad y pesimismo, Joan. Nunca pensé que me alegraría de que le dieran a otra persona la beca para la que yo apliqué, pero gracias a ello hemos podido conocernos. Gracias por apoyarme siempre en buenos y malos momentos. Pero sobre todo gracias por tu sinceridad, incluso cuando se trata de algo que no quiero oir, eso es lo que distingue a un buen amigo.

Quiero dar especialmente las gracias a todos aquellos que han sufrido mi tesis a pesar de estar fuera del laboratorio. Sobre todo quiero acordarme de vosotros, Ana, Rafa, Irene y Victor. Por aguantar horarios extraños. Por entender que desapareciera en épocas de mucho trabajo. Por traerme chocolate al lab cuando no podía más con las revisiones. Por estar siempre ahí.

Bereziki, eskerrak eman nahi dizkiot bide guztian zehar nirekin egon direnei, nere familiari. Egiten dudan edozertan beti laguntzen nauzuelako, eta hartu ditudan erabakietaz fidatu zaretelako, hauen ondorioa ni etxetik urrundu izana izan arren. Gaur naizena eta lortu dudana zuei esker izan da. Nor daki zer datorkigun ondoren. Baina datorrena datorrela, ziur nago edozer gauza gainditu dezakedala zuek alboan izanda.

Y finalmente quiero darte las gracias a ti Borja. Por ser mi compañero en este camino y tener que comprender en ocasiones lo incomprensible. Por creer siempre en mi más que yo misma y no dejar que me rindiera nunca. Por transmitirme tu valentía y tu optimismo. Erronka ugari gainditu ditugu elkarrekin, eta orain berri bat aurkezten zaigu aurrean. Eta ez naiz beldur ez kezkati, edozer gauza gainditzeko gai garela argi daukadalako. Egin dugu lehen eta egingo dugu berriro.

Aquesta tesi ha estat impresa amb el suport de la Fundació IMIM